

# OPTICA ACTA

**Volume 3 1956**

3-5

$dl$ ;  $dl$  = geometrisches Wegelement;  $n = \frac{c}{v}$  = Brechungszahl (des « Mediums ») an der Stelle von  $dl$ . Dementsprechend bezieht sich auch das Prinzip der Konstanz der optischen Weglängen  $\Sigma (n \cdot \Delta l)$  bzw.

$\int_{P_1}^{P_2} n dl$  für zwei oder mehr Strahlen zwischen zwei

Punkten  $P_1$  und  $P_2$  zunächst nur auf benachbarte Strahlen. Bei seiner Anwendung auf « aplanatische » Flächen oder auf « cartesische » Flächen — also Flächen, bei denen nur  $\Sigma (n \cdot \Delta l) = \text{const}$  für alle von einem Punkte ausgehenden Strahlen zwischen diesem Punkte und seinem Bildpunkte gefordert wird, während ja bei den aplanatischen Flächen ausserdem für alle diese Strahlen die Erfüllung der Sinusbedingung erforderlich ist — handelt es sich indessen um eine Vielzahl, eine Mannigfaltigkeit von Strahlen, die keineswegs alle als « einander benachbart » zu bezeichnen sind, wohl aber besitzt hier jeder der betrachteten Strahlen Nachbarstrahlen, mit denen er bzw. seine optische Weglänge verglichen — und als optisch gleich lang verlangt — wird derart, dass diese « Gleichheit » über das ganze Strahlenbündel fortgesetzt werden kann, ohne die Forderung der « Nachbarschaft » bei den einzelnen Vergleichen aufzugeben.

**I. Unterschied der optischen Weglängen eines von einem Achsenpunkt ausgehenden, zur Achse (beliebig) geneigten Strahls und des zugehörigen axialen Strahls zwischen dem gemeinsamen Objektpunkt und ihrem bildseitigen Schnittpunkt bei beliebig grosser sphärischer Aberration.** — Bei der Berechnung der optischen (Licht-) Wege zweier beliebiger, vom gleichen Objektpunkt ausgehender Strahlen von diesem Objektpunkt bis zu ihrem bildseitigen Schnittpunkt bzw. bei der Berechnung ihres optischen Wegunterschiedes handelt es sich aber um diskret liegende Strahlen — z. B. einen oder mehrere (bzw. unendlich viele) Strahlen bestimmter Neigung zu dem « Hauptstrahl » des abbildenden Strahlenbündels und diesen « Hauptstrahl » selbst —, ohne dass hier die Berechtigung der Anwendung des Fermatschen Prinzips oder eines ihm analogen besteht, da es sich jetzt ja nicht — wie bei den aplanatischen oder cartesischen Flächen — um eine kontinuierliche Mannigfaltigkeit von « benachbarten » Strahlen handelt.

Es sei nun zunächst für den Fall der sphärischen Aberration der Unterschied der optischen Weglängen eines zur Achse geneigten und eines mit der Achse zusammenfallenden; vom gleichen Objektpunkt ausgehenden Strahls in ihrem bildseitigen Schnittpunkt berechnet.

Dabei gehen wir aus von einer früheren Arbeit des Verfassers [1], in der die Gleichungen für die Wellenflächen eines Strahlenbündels abgeleitet wurden, die zu einem Strahlenbündel mit sphärischer Aberration gehören.

Die sphärische Aberration, die wir wie üblich durch

$\Delta s'$  bezeichnen, lässt sich (in allen Fällen) darstellen durch die (im allgemeinen endliche) Reihe

$$(1,1) \quad \Delta s' = a \operatorname{tg}^2 u' + b \operatorname{tg}^4 u' + c \operatorname{tg}^6 u' + d \operatorname{tg}^8 u' + \dots,$$

wo  $u'$  der — wie üblich gemessene — Winkel der (bildseitigen) Strahlen des Bündels gegen seine Symmetriachse (Achse des optischen System) sei.

Wie in der angegebenen Arbeit gezeigt wurde, lauten die Gleichungen der zugehörigen Wellenflächen dann — bezogen auf ein Koordinatensystem, dessen  $\xi$ -Achse mit der optischen Achse zusammenfällt,

$$(1,2) \quad \xi = -2a' + a' \operatorname{tg}^2 u' + b' \operatorname{tg}^4 u' + c' \operatorname{tg}^6 u' + \dots - \text{const.} \cos u'$$

$$(1,3) \quad \rho = \frac{2}{1} a' \operatorname{tg} u' + \frac{4}{3} b' \operatorname{tg}^3 u' + \frac{6}{5} c' \operatorname{tg}^5 u' + \dots + \text{const.} \sin u',$$

wenn  $\rho$  den Abstand der Punkte der Wellenflächen von der Achse bedeutet. Hierin ergeben sich die  $a', b', c', \dots$  aus  $a, b, c, \dots$  des  $\Delta s'$  durch die Formeln

$$(1,4) \quad \begin{aligned} a' &= a - \frac{4}{3} b' \\ b' &= b - \frac{6}{5} c' \\ c' &= c - \frac{8}{7} d' \\ d' &= d - \frac{10}{9} e' \\ &\vdots \\ &\vdots \end{aligned}$$

lassen sich also für praktische Fälle — in denen die Reihe für  $\Delta s'$  abbrechen wird — stets leicht berechnen. In Abb. 1 und 2 sind für zwei Strahlenbündel mit sphärischer Aberration diese Aberration, ferner Strahlen, Wellenflächen und rotationssymmetrische sowie entartete Kautikschale dargestellt.

Für unsere Zwecke des Vergleichs der Länge der Lichtwege der die Achse in einem (beliebigen) Punkte treffenden Strahlen mit dem Lichtweg des Achsenstrahls selbst bis zu jenem Punkte benutzen wir keine Voraussetzung über das Abbrechen oder Nichtabbrechen der Darstellung von  $\Delta s'$  (wohl aber natürlich bei einem Nichtabbrechen die Voraussetzung der Konvergenz der Reihe).

Tragen wir vom Achsenschnittpunkt eines der ausseraxialen Strahlen — etwa des bildseitigen Neigungswinkels  $u'_1$  gegen die Achse — eine Strecke (bei  $\Delta s'_1 > 0$  : entgegen der Ausbreitungsrichtung des Lichtes) ab, die gleich dem Betrage  $\Delta s'_1$  der sphärischen Aberration des betreffenden Strahles ist, so gilt für die den betreffenden Strahl dort orthogonal schneidende Wellenfläche nach (1,3)

$$(1,5) \quad \frac{\rho}{\sin u'_1} = \Delta s'_1 = \frac{1}{\cos u'_1} (2a' + \frac{4}{3} b' \operatorname{tg}^2 u'_1 + \frac{6}{5} c' \operatorname{tg}^4 u'_1 + \dots) + \text{const.},$$



so dass die zu dieser Wellenfläche gehörige Konstante den Wert hat

$$(1,6) \quad \text{const}_1 = \Delta s'_1 - \frac{1}{\cos u'_1} (2 a' + \frac{4}{3} b' \text{tg}^2 u'_1 + \frac{6}{5} c' \text{tg}^4 u'_1 + \dots) .$$

Mit dieser Konstanten lauten die Gleichungen der betreffenden Wellenfläche

$$\begin{aligned} \xi_1 = & -2 a' \left( 1 - \frac{\cos u'}{\cos u'_1} \right) + \\ & + \left( a' \text{tg}^2 u' + \frac{4}{3} b' \frac{\text{tg}^2 u'_1}{\cos u'_1} \cos u' \right) + \\ & + \left( b' \text{tg}^4 u' + \frac{6}{5} c' \frac{\text{tg}^4 u'_1}{\cos u'_1} \cos u' \right) + \\ & + \dots - \Delta s'_1 \cdot \cos u' \end{aligned}$$

$\xi$ - und  $\rho$ -Werte der einzelnen Punkte der betreffenden speziellen Wellenfläche liefert, die vom Achsen-schnittpunkt des Strahls  $u'_1$  längs dieses Strahls den Abstand  $-\Delta s'_1$  besitzt.

Für  $u' = 0$  wird  $(\rho_1)_0 = 0$  — wie erforderlich — und

$$(\xi_1)_0 = -2 a' \left( 1 - \frac{1}{\cos u'_1} \right) + \frac{4}{3} b' \frac{\text{tg}^2 u'_1}{\cos u'_1} + \frac{6}{5} c' \frac{\text{tg}^4 u'_1}{\cos u'_1} + \dots - \Delta s'_1 .$$

Setzen wir hier für  $\Delta s'_1$  den Wert aus (1,5) ein, so wird für  $u' = 0$

$$(1,7) \quad (\xi_1)_0 = -2 a' - \text{const}_1 .$$

Da andererseits für den Strahl mit  $u'_1 = 0$  auch  $\Delta s'_1 = 0$  ist und die entsprechend wie oben konstruierte Wellenfläche durch die Kaustikspitze geht,

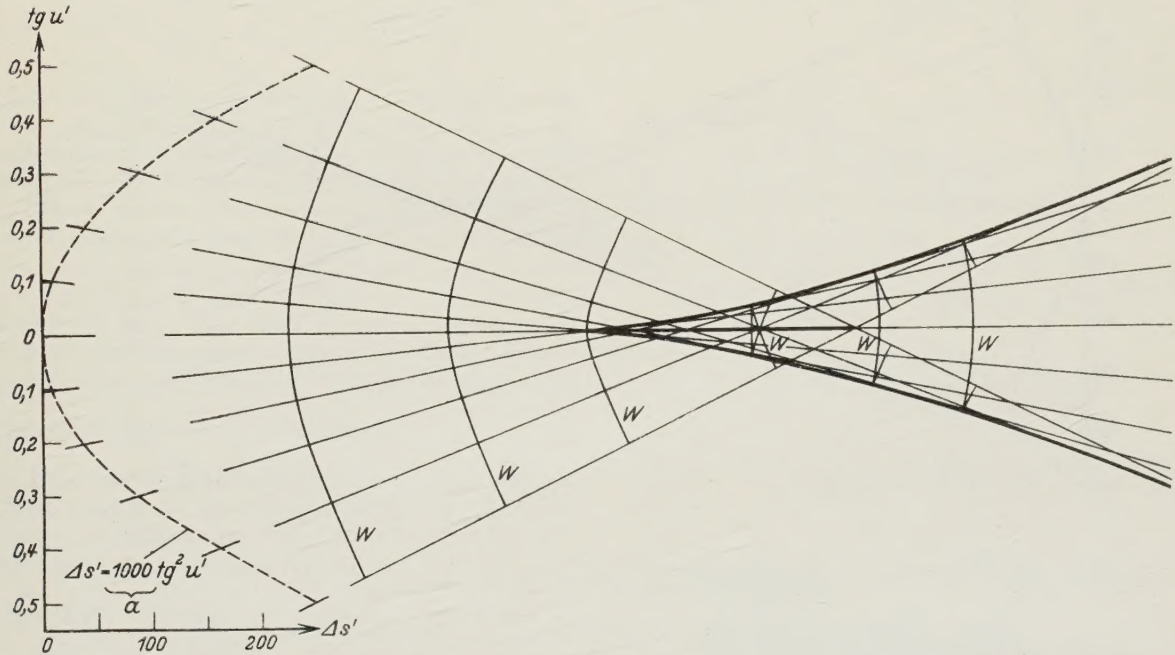


ABB. 1. — Strahlenaberration, Strahlenverlauf, Wellenflächen und Kaustik eines mit sphärischer Aberration behafteten Strahlenbündels für den Fall, dass

$$\Delta s' = a \text{tg}^2 u'$$

$$\text{mit } a = 1000 \left( \text{bzw. } \frac{1}{6} \right)$$

$$\begin{aligned} \rho_1 = & 2 a' \left( \text{tg} u' - \frac{\sin u'}{\cos u'_1} \right) + \\ & + \frac{4}{3} b' \left( \text{tg}^3 u' - \frac{\text{tg}^2 u'_1}{\cos u'_1} \sin u' \right) + \\ & + \frac{6}{5} c' \left( \text{tg}^5 u' - \frac{\text{tg}^4 u'_1}{\cos u'_1} \sin u' \right) + \dots + \Delta s'_1 \cdot \sin u' , \end{aligned}$$

worin  $u'$  wieder der variable Parameter ist, der die

so gilt für diese Wellenfläche nach (1,6)  $\text{const}_0 = -2 a'$  und  $(\xi_0)_0 = 0$ , wobei also der Koordinaten-Ursprungspunkt in die Spitze der Kaustik gelegt ist.

Sollte also die Wellenfläche, die den Strahl der Achsenneigung  $\text{tg} u'_1$  im Abstand  $-\Delta s'_1$  von seinem Achsen-schnittpunkt trifft, durch die Spitze der Kaustik gehen, sollten also — anders ausgedrückt — Achsenstrahl und Strahl der Neigung  $\text{tg} u'_1$  in ihrem Schnittpunkt gleiche Lichtweglängen vom gemeinsamen, auf der Achse liegenden Objektpunkt aus zurückgelegt

haben, so müsste auch  $\text{const}_1 = -2 a'$  sein, was indessen nach (1,6) nicht der Fall ist. Der Unterschied der beiden Konstanten ist demnach der Lichtwegunterschied, den der achsengeneigte Strahl gegen den Achsenstrahl in ihrem gemeinsamen Schnittpunkt besitzen. Bezeichnen wir diesen Wegunterschied durch

$$\Delta w_{1/0} = w(\text{tg } u'_1) - w_0,$$

so ist  $\Delta w_{1/0} = (\xi_1)_0 - (\xi_0)_0 = (\xi_1)_0$ , also

$$\Delta w_{1/0} = -2 a' - \Delta s'_1 + \frac{1}{\cos u'_1} \left( 2 a' + \frac{4}{3} b' \text{tg}^2 u'_1 + \frac{6}{5} c' \text{tg}^4 u'_1 + \dots \right)$$

bei sphärischer Aberration, nur gilt, wenn  $u'_1$  genügend klein ist, so dass man beide Strahlen als (unendlich) benachbart bezeichnen kann. Denn (nur) für diesen Fall wird ja  $\Delta w_{1/0} = 0$ .

II. — Gleichungen des Meridianschnitts der Wellenflächen eines komabehafteten Strahlenbündels und Unterschied der optischen Weglängen eines zum Hauptstrahl geneigten Strahls der Meridianebene und des Hauptstrahls zwischen dem gemeinsamen Objektpunkt und ihrem bildseitigen Schnittpunkt. — Entsprechend zu der Darstellung der sphärischen Aberration in ihrer Abhängigkeit vom Tangens des bildseitigen Neigungswinkels  $u'$  der einzelnen Strahlen des

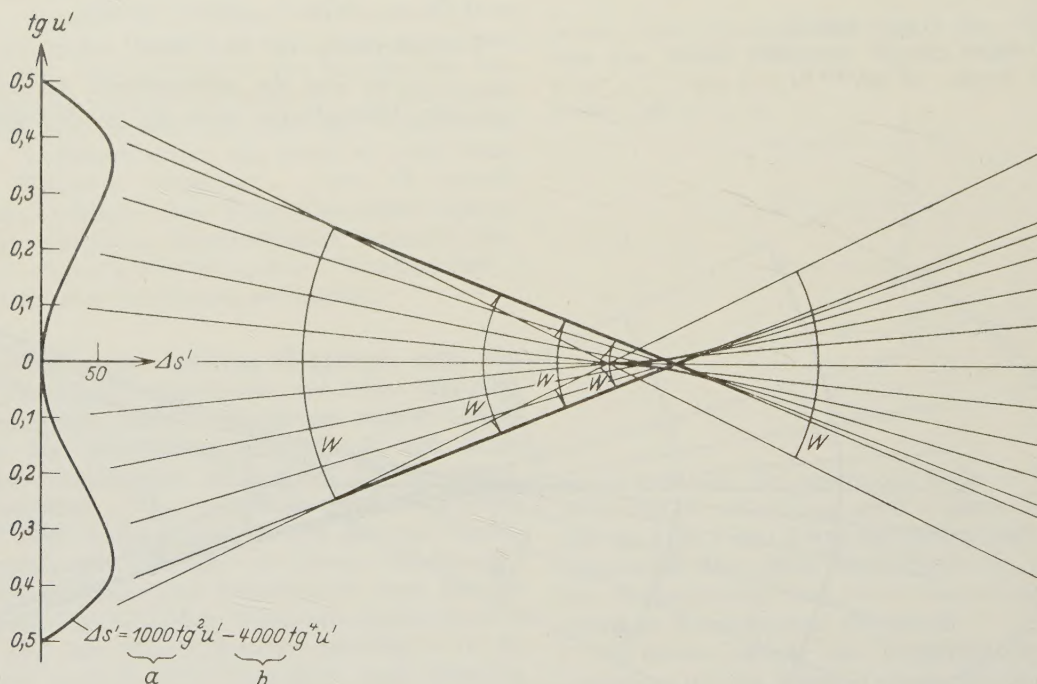


ABB. 2. — Strahlenaberration, Strahlenverlauf, Wellenflächen und Kaustik eines mit sphärischer Aberration behafteten Strahlenbündels für den Fall, dass

$$\Delta s' = a \text{tg}^2 u' + b \text{tg}^4 u' \text{ mit } a = 1000 \left( \text{bzw. } \frac{1}{6} \right), \quad b = -4000 \left( \text{bzw. } -\frac{2}{3} \right)$$

$$= \left( \frac{1}{\cos u'_1} - 1 \right) \left( 2 a' + \Delta s'_1 \right) - \frac{\text{tg}^2 u'_1}{\cos u'_1} \times (a' + b' \text{tg}^2 u'_1 + c' \text{tg}^4 u'_1 + \dots).$$

Aus diesem Ausdruck ersieht man also auch unmittelbar, dass der Satz von der konstanten Lichtweglänge zweier vom gleichen Objektpunkt ausgehender Strahlen von diesem Objektpunkt bis zu ihrem gemeinsamen Schnittpunkt nicht in dieser Allgemeinheit gilt, auch nicht, wenn der eine der beiden Strahlen der Achsenstrahl ist, sondern dass dieser Satz auch hier,

abbildenden Strahlenbündels zur Symmetrieachse des abbildenden Systems lässt sich auch der Komafehler als Potenzreihe der Neigungswinkel  $\kappa'$  der Strahlen des Strahlenbündels gegen dessen Schwerstrahl (Bündelachse) darstellen, sofern wir uns auf Strahlen des Meridianschnitts beschränken.

Bezeichnen wir — als Analogon zu  $\Delta s'$  — den Abstand des Schnittpunktes eines der im Meridianschnitt verlaufenden Strahlen mit dem Schwerstrahl vom Schnittpunkt des Schwerstrahls mit einem ihm unmittelbar benachbart verlaufenden Strahl des Meridianschnitts durch  $\Delta l$ , so lässt sich  $\Delta l$  stets als endliche



oder unendliche Potenzreihe von  $\operatorname{tg} \kappa'$  darstellen :

$$(2,1) \quad \Delta l' = a_1 \operatorname{tg} \kappa' + a_2 \operatorname{tg}^2 \kappa' + a_3 \operatorname{tg}^3 \kappa' + \dots + a^4 \operatorname{tg}^4 \kappa' + \dots$$

oder, mit  $\operatorname{tg} \kappa' = \frac{1}{y}$ ,

$$(2,2) \quad \Delta l' = \frac{a_1}{y} + \frac{a_2}{y^2} + \frac{a_3}{y^3} + \frac{a_4}{y^4} + \dots = \varphi\left(\frac{1}{y}\right).$$

Die von MERTÉ (W. MERTÉ, *ZS. f. Phys.* **33**, 1925, 533) zur Berechnung der Wellenflächen bei sphärischer Aberration des betreffenden Strahlenbündels angegebenen Formeln

$$(2,3) \quad \left\{ \begin{array}{l} \xi = \frac{y}{\sqrt{1+y^2}} \int_{\text{const}}^y \frac{\Phi(y)}{\sqrt{1+y^2}} dy \\ \rho = -\frac{1}{\sqrt{1+y^2}} \int_{\text{const}}^y \frac{\Phi(y)}{\sqrt{1+y^2}} dy + \frac{1}{y} \varphi\left(\frac{1}{y}\right) \end{array} \right.$$

mit

$$(2,4) \quad \Phi(y) = \frac{1}{y} \frac{d}{dy} \varphi\left(\frac{1}{y}\right) - \frac{1}{y^2} \varphi\left(\frac{1}{y}\right)$$

behalten auch für den Komafehler — wenigstens für die uns zunächst interessierenden Verhältnisse im Meridianschnitt — ihre Gültigkeit, wenn wir jetzt die  $\xi$ -Achse mit dem Hauptstrahl des Bündels zusammenfallen lassen und die  $\rho$ -Achse senkrecht zur  $\xi$ -Achse in der Meridianebene liegt. Man überzeugt sich leicht davon, wenn man aus (2,3) die Tangente an die Kurve ( $\xi, \rho$ ) in irgend einem Punkte ( $y$ ) der Kurve berechnet. Man erkennt dann, dass diese Tangente senkrecht zu dem Strahl des (ebenen) Strahlenbündels liegt, dessen Neigung gegen den Schwerstrahl (Hauptstrahl) des Bündels durch  $\operatorname{tg} \kappa' = \frac{1}{y}$  bestimmt ist.

Zur Berechnung der Wellenflächen — oder besser : der Wellenlinien (in der Meridianebene) des komabehafteten Strahlenbündels bilden wir nach (2,4) zunächst

$$(2,5) \quad \Phi(y) = -2 \frac{a_1}{y^3} - 3 \frac{a_2}{y^4} - 4 \frac{a_3}{y^5} - \dots - 5 \frac{a_4}{y^6} - \dots$$

Damit haben wir nach (2,3) die Integrale zu berechnen

$$\begin{aligned} \int_{\text{const}}^y \frac{\Phi(y)}{\sqrt{1+y^2}} dy &= -2 a_1 \int \frac{dy}{y^3 \sqrt{1+y^2}} \\ &- 3 a_2 \int \frac{dy}{y^4 \sqrt{1+y^2}} - 4 a_3 \int \frac{dy}{y^5 \sqrt{1+y^2}} - \dots \\ &= - \sum_{n=1}^{\infty} 2 n a_{2n-1} \int \frac{dy}{y^{2n+1} \sqrt{1+y^2}} - \\ &- \sum_{n=1}^{\infty} (2n+1) a_{2n} \int \frac{dy}{y^{2n+2} \sqrt{1+y^2}}. \end{aligned}$$

Wir bezeichnen die in der ersten  $\Sigma$  auftretenden Integrale durch  $I_{2n+1}$ , die in der zweiten  $\Sigma$  auftretenden Integrale durch  $I_{2(n+1)}$ . Bei den  $I_{2(n+1)}$  handelt es sich um die gleichen Integrale, die auch bei der Berechnung der Wellenflächen eines mit sphärischer Aberration behafteten Strahlenbündels auftreten. Für sie gilt

$$-(2n+1) I_{2(n+1)} = \frac{\sqrt{1+y^2}}{y^{2n+1}} \sum_0^n (-1)^\mu \frac{\binom{n}{\mu}}{\binom{n-0,5}{\mu}} y^{2\mu}.$$

Entsprechend findet man durch wiederholte partielle Integration und dadurch gewonnene Zurückführung der Integrale auf niedriger indizierte

$$\begin{aligned} -2n I_{2n+1} &= \\ &= (-1)^n \frac{2n}{2^{2n}} \binom{2n}{n} \Re \Im \sin \frac{1}{y} \\ &- 2n \binom{2n}{n} \sum_{\mu=1}^n (-1)^\mu \frac{1}{4^\mu \binom{2(n-\mu)}{n-\mu} \binom{2(n-\mu)+1}{n-\mu}} \times \\ &\quad \times \frac{\sqrt{1+y^2}}{y^{2(n+1-\mu)}}. \end{aligned}$$

Wir erhalten so \*

$$\begin{aligned} \int \frac{\Phi(y)}{\sqrt{1+y^2}} dy &= \\ &= \sum_{n=1}^{\infty} a_{2n} \frac{\sqrt{1+y^2}}{y^{2n+1}} \left\{ \sum_{\mu=0}^n (-1)^\mu \frac{\binom{n}{\mu}}{\binom{n-0,5}{\mu}} y^{2\mu} \right\} \\ &- \sum_{n=1}^{\infty} a_{2n-1} \frac{\sqrt{1+y^2}}{y^{2(n+1)}} 2n \binom{2n}{n} \\ &\left\{ \sum_{\mu=1}^n (-1)^\mu \frac{1}{4^\mu \binom{2(n-\mu)}{n-\mu} \binom{2(n-\mu)+1}{n-\mu}} y^{2\mu} \right\} + \\ &+ \left( \Re \Im \sin \frac{1}{y} \right) \sum_{n=1}^{\infty} (-1)^n a_{2n-1} \frac{n}{2^{2n-1}} \binom{2n}{n}. \end{aligned}$$

Schreiben wir die hier auftretenden Summen ausführlich, so lassen sich wieder — wie bei der Behandlung der sphärischen Aberration — gewisse Zusammenfassungen durch Einführung neuer Koeffizienten durchführen.

Wir erhalten

$$(2,6) \quad \int \frac{\Phi(y)}{\sqrt{1+y^2}} dy =$$

$$(*) \text{ Statt } \frac{\binom{n}{\mu}}{\binom{n-0,5}{\mu}} \text{ lässt sich auch schreiben :}$$

$$2^{2\mu} \frac{\binom{n}{\mu} \binom{n-1}{\mu}}{\binom{2n-1}{2\mu} \binom{2\mu}{\mu}}$$

$$= \frac{\sqrt{1+y^2}}{y} \left( -2 B_2 + \frac{B_2}{y^2} + \frac{B_4}{y^4} + \frac{B_6}{y^6} + \frac{B_8}{y^8} + \dots \right) \\ + \frac{\sqrt{1+y^2}}{y^2} \left( B_1 + \frac{B_3}{y^2} + \frac{B_5}{y^4} + \frac{B_7}{y^6} + \frac{B_9}{y^8} + \dots \right) \\ - B_1 \mathfrak{N}r \sin \frac{1}{y}.$$

Die hier auftretenden Koeffizienten  $B_i$  (mit  $i = 1, 2, \dots$ ) ergeben sich aus den Koeffizienten der Reihenentwicklung (2,1) oder (2,2) von  $\Delta l'$  nach folgenden Rekursionsformeln:

$$(2,7) \quad \left\{ \begin{array}{ll} B_2 = a_2 - \frac{4}{3} B_4 & B_1 = a_1 - \frac{3}{2} B_3 \\ B_4 = a_4 - \frac{6}{5} B_6 & B_3 = a_3 - \frac{5}{4} B_5 \\ B_6 = a_6 - \frac{8}{7} B_8 & B_5 = a_5 - \frac{7}{6} B_7 \\ B_8 = a_8 - \frac{10}{9} B_{10} & B_7 = a_7 - \frac{9}{8} B_9 \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \end{array} \right.$$

Bricht die Reihenentwicklung (2,1) oder (2,2) von  $\Delta l'$  ab, so brechen auch die Koeffizientenreihen der  $B_i$  automatisch ab.

Aus (2,6) und (2,3) ergeben sich so die Koordinaten der Schnittpunkte der Meridianebene mit den Wellenflächen des komabehafteten Strahlenbündels:

$$(2,8) \quad \left\{ \begin{array}{l} \xi = -2 B_2 + (B_2 \operatorname{tg}^2 x' + B_4 \operatorname{tg}^4 x' + \\ \quad + B_6 \operatorname{tg}^6 x' + B_8 \operatorname{tg}^8 x' + \dots \\ \quad + (B_1 \operatorname{tg} x' + B_3 \operatorname{tg}^3 x' + B_5 \operatorname{tg}^5 x' + \\ \quad + B_7 \operatorname{tg}^7 x' + \dots) \\ \quad - [B_1 \mathfrak{N}r \sin (\operatorname{tg} x') - \text{const}] \cos x' \\ \eta = \left( 2 B_2 \operatorname{tg} x' + \frac{4}{3} B_4 \operatorname{tg}^3 x' + \right. \\ \quad + \frac{6}{5} B_6 \operatorname{tg}^5 x' + \frac{8}{7} B_8 \operatorname{tg}^7 x' + \dots \Big) \\ \quad + \left[ \frac{3}{2} B_3 \operatorname{tg}^2 x' + \frac{5}{4} B_5 \operatorname{tg}^4 x' + \right. \\ \quad \left. + \frac{7}{6} B_7 \operatorname{tg}^6 x' + \dots \right] \\ \quad + [B_1 \mathfrak{N}r \sin (\operatorname{tg} x') - \text{const}] \sin x' \end{array} \right.$$

worin noch — wir bekannt —

$$\mathfrak{N}r \sin (\operatorname{tg} x') = \mathfrak{N}r \sin \frac{1}{y} = \\ = \sum_{\mu=0}^{\infty} (-1)^{\mu} \frac{(2\mu)!}{2^{\mu}\mu!} \frac{1}{(2\mu+1)y^{2\mu+1}}.$$

Abbildungen 3, 4 und 5 geben eine Darstellung der Strahlen und Wellenfläche eines komabehafteten Strahlenbündels für 3 verschiedene Fälle.

Untersuchen wir noch die Länge der Lichtwege eines beliebigen der Strahlen der Meridianebene und des

Schwerstrahls (Hauptstrahls) vom Objektpunkt bis zu ihrem gemeinsamen Schnittpunkt im Bildraum.

Zu dem Zwecke greifen wir wieder einen beliebigen, dann aber bestimmten der Strahlen heraus. Sein Neigungswinkel gegen den Hauptstrahl sei  $x'_1$ . Auf ihm tragen wir vom Schnittpunkt mit dem Hauptstrahl die zugehörige Länge —  $\Delta l'_1$  ab und betrachten die durch den Endpunkt der abgetragenen Strecke —  $\Delta l'_1$  gehende Wellenfläche bzw. deren Schnittpunkt mit der Meridianebene. Für diese gilt

$$(2,10) \quad \frac{\eta_1}{\sin x'_1} = \Delta l'_1$$

so dass für diese Wellenfläche die Konstante nach (2,9) den Wert

$$(2,11) \quad \text{const}_1 = -\Delta l'_1 + \\ + \frac{1}{\cos x'_1} \left( 2 B_2 + \frac{4}{3} B_4 \operatorname{tg}^2 x'_1 + \frac{6}{5} B_6 \operatorname{tg}^4 x'_1 + \dots \right) \\ + \frac{1}{\cos x'_1} \left[ \frac{3}{2} B_3 \operatorname{tg} x'_1 + \frac{5}{4} B_5 \operatorname{tg}^3 x'_1 + \dots \right] \\ + B_1 \mathfrak{N}r \sin (\operatorname{tg} x'_1)$$

Die Gleichungen dieser betreffenden Wellenfläche lauten mit dieser Konstanten (2,11) nach (2,8) bzw. (2,9) demnach

$$\xi_1 = -2 B_2 \left( -1 \frac{\cos x'_1}{\cos x'_1} \right) + \\ + \left[ \left( B_2 \operatorname{tg}^2 x'_1 + \frac{4}{3} B_4 \frac{\cos x'_1}{\cos x'_1} \operatorname{tg}^2 x'_1 \right) + \right. \\ \left. + \left( B_4 \operatorname{tg}^4 x'_1 + \frac{6}{5} B_6 \frac{\cos x'_1}{\cos x'_1} \operatorname{tg}^4 x'_1 \right) + \dots \right] \\ + \left[ \left( B_1 \operatorname{tg} x'_1 + \frac{3}{2} B_3 \frac{\cos x'_1}{\cos x'_1} \operatorname{tg} x'_1 \right) + \right. \\ \left. + \left( B_3 \operatorname{tg}^3 x'_1 + \frac{5}{4} B_5 \frac{\cos x'_1}{\cos x'_1} \operatorname{tg}^3 x'_1 \right) + \dots \right] \\ + [-B_1 \langle \mathfrak{N}r \sin (\operatorname{tg} x'_1) - \mathfrak{N}r \sin (\operatorname{tg} x'_1) \rangle - \Delta l'_1] \cos x'_1; \\ \eta_1 = \left[ 2 B_2 \left( \operatorname{tg} x'_1 - \frac{\sin x'_1}{\cos x'_1} \right) + \right. \\ \left. + \frac{4}{3} B_4 \left( \operatorname{tg}^3 x'_1 - \frac{\sin x'_1}{\cos x'_1} \operatorname{tg}^2 x'_1 \right) \right. \\ \left. + \frac{6}{5} B_6 \left( \operatorname{tg}^5 x'_1 - \frac{\sin x'_1}{\cos x'_1} \operatorname{tg}^4 x'_1 \right) + \dots \right] \\ + \left[ \frac{3}{2} B_3 \left( \operatorname{tg}^2 x'_1 - \frac{\sin x'_1}{\cos x'_1} \operatorname{tg} x'_1 \right) \right. \\ \left. + \frac{5}{4} B_5 \left( \operatorname{tg}^4 x'_1 - \frac{\sin x'_1}{\cos x'_1} \operatorname{tg}^3 x'_1 \right) + \dots \right] \\ + [B_1 \langle \mathfrak{N}r \sin (\operatorname{tg} x'_1) - \mathfrak{N}r \sin (\operatorname{tg} x'_1) \rangle + \Delta l'_1] \sin x'_1.$$



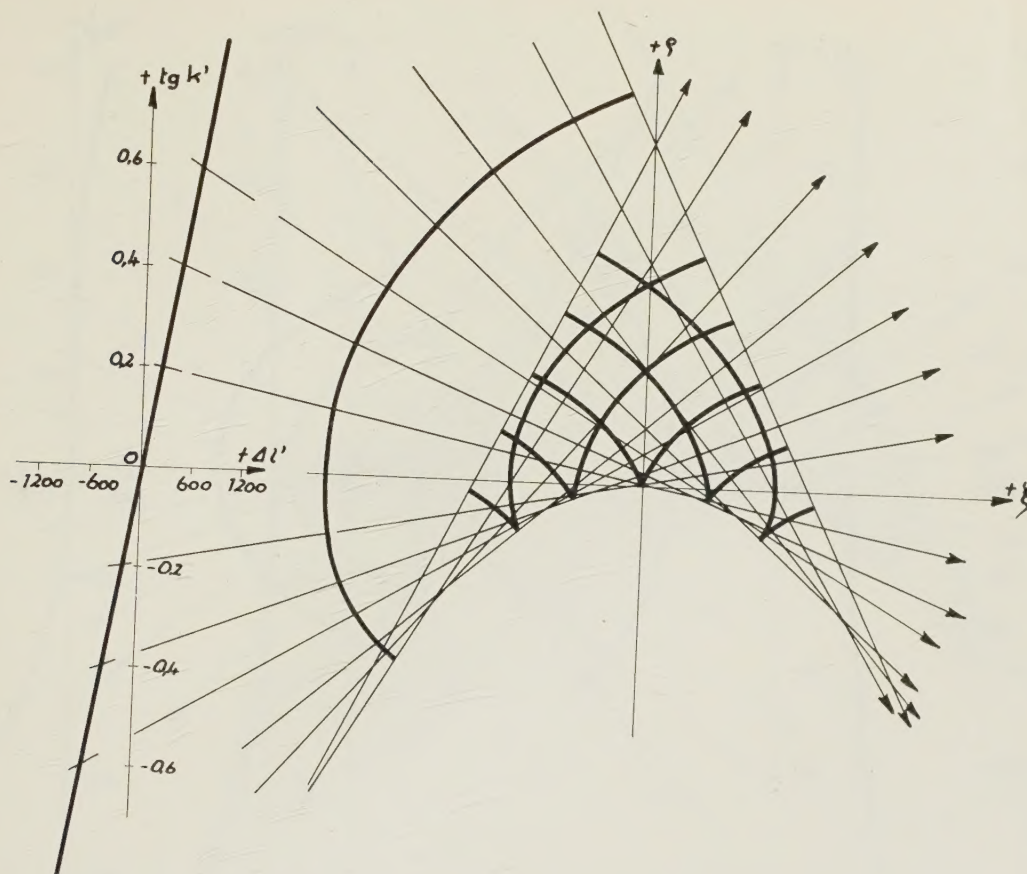


ABB. 3. — Strahlenaberration, Strahlenverlauf und Wellenflächen [im Schnitt mit der Meridianebene des bildseitigen Hauptstrahls (= Schwerstrahls)] eines mit Koma behafteten Strahlenbündels für den Fall, dass

$$\Delta l' = a_1 \operatorname{tg} \kappa'; a_1 = \frac{1}{6} \text{ bzw. } 1\,000.$$

Für  $\kappa' = 0$  wird  $\gamma_1 = (\gamma_1)_0 = 0$  — wie erforderlich — und

$$\xi_1 = (\xi_1)_0 = -2 B_2 \left( 1 - \frac{1}{\cos \kappa'_1} \right) + \frac{1}{\cos \kappa'_1} \operatorname{tg}^2 \kappa'_1 \times$$

$$\begin{aligned} & \left[ \frac{4}{3} B_4 + \frac{6}{5} B_6 \operatorname{tg}^2 \kappa'_1 + \right. \\ & \quad \left. + \frac{8}{7} B_8 \operatorname{tg}^4 \kappa'_1 + \dots \right] \\ & + \frac{1}{\cos \kappa'_1} \operatorname{tg} \kappa'_1 \left[ \frac{3}{2} B_3 + \frac{5}{4} B_5 \operatorname{tg}^2 \kappa'_1 + \right. \\ & \quad \left. + \frac{7}{6} B_7 \operatorname{tg}^4 \kappa'_1 + \dots \right] \\ & + B_1 \operatorname{tg} \sin (\operatorname{tg} \kappa'_1) - \Delta l'_1. \end{aligned}$$

Setzen wir hier wieder den Wert für  $\Delta l'_1$  aus (2,11) ein, so ergibt sich für  $\kappa' = 0$ , also für den Schnittpunkt der betreffenden, durch  $\kappa'_1$  und  $\Delta l'_1$  bestimmten Wellenfläche mit dem Hauptstrahl des Bündels

$$(\xi_1)_0 = -2 B_2 + \operatorname{const}_1.$$

Für einen dem Schwerstrahl (Hauptstrahl) unmittelbar benachbarten Strahl gilt  $\kappa'_1 \rightarrow 0$  und  $\Delta l'_1 \rightarrow 0$ , so dass für einen solchen Strahl die Konstante ( $\operatorname{const}_0$ ) der nach der gleichen Vorschrift wie oben konstruierten Wellenfläche nach (2,11) den Wert

$$\operatorname{const}_0 = +2 B_2$$

annimmt, also  $(\xi_0)_0 = 0$  wird, der Ursprungspunkt des  $(\xi, \rho)$  — Systems also im Schnittpunkt des Haupt-

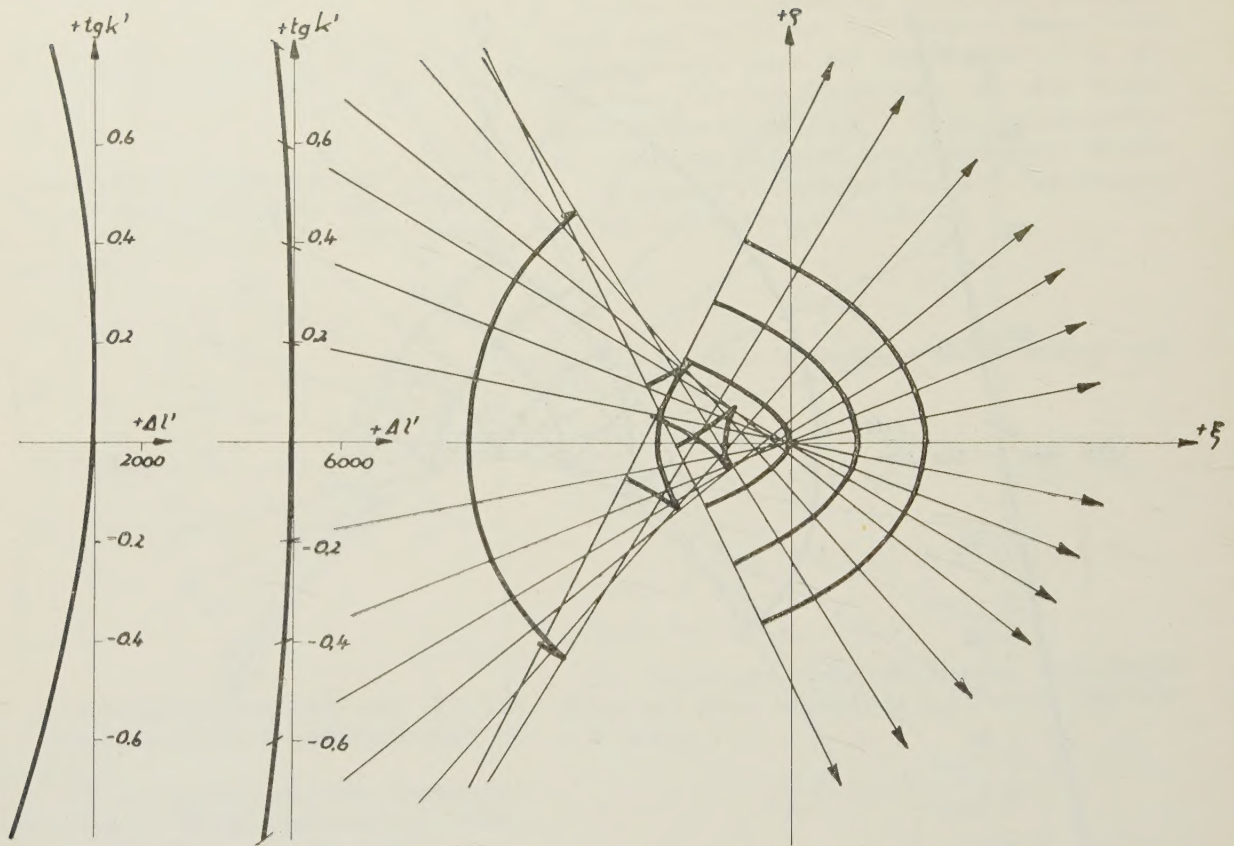


ABB. 4. — Strahlenaberration, Strahlenverlauf und Wellenflächen (im Schnitt mit der Meridianebene des bildseitigen Hauptstrahls) eines mit Koma behafteten Strahlenbündels für den Fall, dass

$$\Delta l' = a_1 \operatorname{tg} \kappa' + a_2 \operatorname{tg}^2 \kappa'; a_1 = \frac{1}{6} \text{ bzw. } 1\,000; a_2 = -\frac{2}{3} \text{ bzw. } -4\,000.$$

strahls des komabehafteten Strahlenbündels mit den ihm unendlich benachbarten Strahlen liegt.

Wäre  $\operatorname{const}_1 = \operatorname{const}_0$ , so würde es sich bei beiden Wellenflächen um die gleiche Wellenfläche handeln, der zum Hauptstrahl geneigte Strahl (mit  $\Delta l' = \Delta l'_1$ ) sowie der Hauptstrahl selbst hätten dann also in ihrem gemeinschaftlichen Schnittpunkt gleichen Lichtwegabstand vom gemeinschaftlichen Objektpunkt. Tatsächlich aber ist  $\operatorname{const}_1 \neq \operatorname{const}_0$ .

Der Wertunterschied beider Konstanten gibt wieder unmittelbar den Lichtwegunterschied beider Strahlen von dem beiden gemeinsamen Objektpunkt bis zu ihrem gemeinschaftlichen Schnittpunkt im Bildraum. Wir bezeichnen diesen Wegunterschied wieder durch

$$\Delta W_{1/0} = W(\operatorname{tg} \kappa'_1) - W_0 = (\xi_1)_0 - (\xi_0)_0 = (\xi_1)_0.$$

Dann ist nach (2,11)

$$\Delta W_{1/0} = -2 B_2 - \Delta l'_1 + \frac{1}{\cos \kappa'_1} \left( 2 B_2 + \frac{4}{3} B_4 \operatorname{tg}^2 \kappa'_1 + \frac{6}{5} B_6 \operatorname{tg}^4 \kappa'_1 + \dots \right)$$

$$+ \frac{1}{\cos \kappa'_1} \left[ \frac{3}{2} B_3 \operatorname{tg} \kappa'_1 + \frac{5}{4} B_5 \operatorname{tg}^3 \kappa'_1 + \dots \right]$$

$$+ B_1 \Re \sin(\operatorname{tg} \kappa'_1).$$

$$\Delta W_{1/0} = \left( \frac{1}{\cos \kappa'_1} - 1 \right) (2 B_2 + \Delta l'_1)$$

$$- \frac{\operatorname{tg} \kappa'_1}{\cos \kappa'_1} (B_1 + B_2 \operatorname{tg} \kappa'_1 + B_3 \operatorname{tg}^2 \kappa'_1 + B_4 \operatorname{tg}^3 \kappa'_1 + B_5 \operatorname{tg}^4 \kappa'_1 + \dots) + B_1 \Re \sin(\operatorname{tg} \kappa'_1).$$



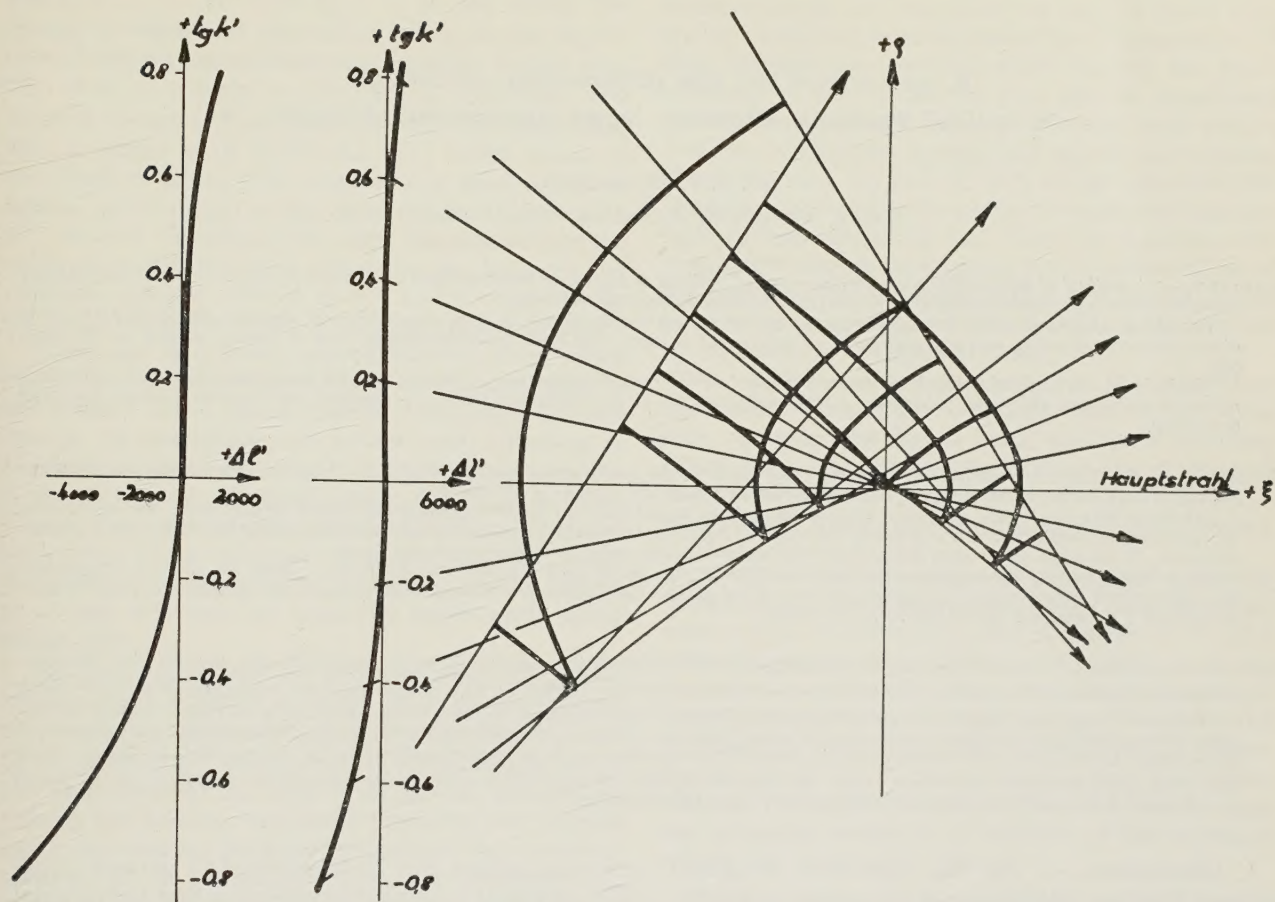


ABB. 5. — Strahlenaberration, Strahlenverlauf und Wellenflächen (im Schnitt mit der Meridianebene des bildseitigen Hauptstrahls) eines mit Koma behafteten Strahlenbündels für den Fall, dass

$$\Delta l' = a_1 \operatorname{tg} \kappa' + a_2 \operatorname{tg}^2 \kappa' + a_3 \operatorname{tg}^3 \kappa'; \quad a_1 = \frac{1}{6} \text{ bzw. } 1\,000; \quad a_2 = -\frac{2}{3} \text{ bzw. } -4\,000; \quad a_3 = 1 \text{ bzw. } 6\,000.$$

**Schlussbemerkung.** — Bei den vorstehenden Untersuchungen, die für die Beurteilung der Korrektur eines optischen Systems von Bedeutung sind [2], handelt es sich allein um die Bestimmung der Lichtwegunterschiede. Die in dem Gebiet der Brennflächen bzw. beim Durchgang durch die beiden Kaustikschalen — von denen die eine evtl. entartet ist — auftretenden Phasensprünge sind hier nicht berücksichtigt.

#### LITERATUR

- [1] J. PICT, *ZS. f. Instrkde* **51**, 1931, 77-81.  
 [2] Siehe das inzwischen (Dezember 1955) im Verlag der Wissenschaften, Berlin erschienene Buch des Verfassers: „Grundlagen der geometrisch-optischen Abbildung“.

*Manuscript reçu le 12 novembre 1955.*



## A procedure for the differential correction of optical systems allowing large parameter changes.

W. S. S. BLASCHKE.

Cooke, Troughton & Simms, Ltd., York. \*

**SUMMARY.** — A method of differential correction for optical systems is developed which serves in the coarser analysis of the aberrational characteristics of the system as well as in the final detailed balance of residuals.

Variations of optical paths through a general optical system due to relatively large constructional changes are considered, both at the changed refracting surfaces and surfaces subsequent to them. The extra-axial meridional case is treated as well as the axial case.

The technique is applied to two high numerical aperture microscope objectives suffering from the marginal and zonal unbalance normally associated with such objectives. A substantial change in the wavefront shapes is achieved, resulting in a much improved correction.

**SOMMAIRE.** — On étudie une méthode de correction différentielle des systèmes optiques utilisable pour une analyse grossière des aberrations d'un système aussi bien que pour la correction finale des résidus d'aberrations.

Les variations de chemin optique à travers un système ayant subi des modifications relativement importantes, des paramètres de construction sont étudiées à la fois sur les surfaces réfringentes modifiées et sur les surfaces suivantes. On étudie dans le plan méridien, le cas d'un point objet éloigné de l'axe aussi bien que celui de l'objet situé sur l'axe.

La méthode est appliquée à deux objectifs de microscope de grande ouverture numérique, affectés de défauts d'équilibre des aberrations marginales et zonales propres à ces sortes d'objectifs. Un changement notable des formes du front d'onde est calculé ; il améliore sensiblement la correction.

**ZUSAMMENFASSUNG.** — Es wird eine Differentialmethode für die Korrektur optischer Systeme entwickelt, die sich bei der Probenanalyse des Systems ebenso gut bewährt wie bei der Feinkorrektur der Restfehler. Die Änderungen der optischen Weglänge durch ein allgemeines optisches System bei verhältnismässig starken Änderungen im Aufbau, werden untersucht und zwar sowohl an der geänderten brechenden Fläche als auch an den anschliessenden Flächen. Der Meridionalstrahl wird dabei ebenso gut erfasst wie der achsiale Strahl. Dieses Verfahren wird auf zwei Mikroskopobjektive hoher numerischer Apertur angewendet, bei denen Rand und Zone, wie bei solchen Objektiven üblich, nicht ganz ausgeglichen sind. Man erreicht eine wesentliche Änderung in der Form der Wellenfront und damit einen sehr viel besseren Korrektionszustand.

**I. Introduction.** — The fine correction of optical systems based on calculations of increments in aberrations with small constructional changes is today a well established procedure of optical design. The technique has been presented in various forms (CRUIKSHANK & MC AULEY, 1945-46; FREITAG, 1952; ROSEN & ELDERT, 1954; LUCY, 1955), and has been adapted, by the latter three authors for automatic computation. Throughout all these papers the difficulty presented by the non-linear dependance of the aberration on the parameters and the limitation of the technique to very small changes is accepted.

The problem has been stated recently in a different way by HOPKINS (unpublished) as follows. If a small change is made in the construction of the system, say a change in the curvature of a surface, the wavefront incident upon this surface will travel in the modified system, at a particular zone, a greater distance in one medium, and a correspondingly smaller distance in the adjacent medium, than before. This introduces a change in the departure of the wavefront from a reference sphere and hence a change in aberration. If however the parameter change is a first order small quantity, the path taken by the ray in the modified system will be neighbouring to that taken by the original ray. Ray paths will be equal by FERMAT'S Theorem and the change in wavefront departure from

reference sphere may be transferred to the exit pupil. The argument is completed by noting that the paraxial focus has moved and that a change in reference sphere to the new focus is needed to obtain the required value of the aberration change. This approach leads to much simpler equations than those in the above papers, while making exactly the same approximations. Moreover it lays bare more precisely the nature of approximations made and shows how the technique may be modified to give the aberrational increments for very much larger constructional changes. It is the purpose of the present paper to develop this extension and thus provide a method which will assist in the coarser analysis of an optical system as well as in its final correction.

**II. Treatment of Approximations.** — Let  $\Sigma$  be an optical system with ray  $O\Pi$  (wavefront  $\bar{o}\Pi$ ) incident upon it. At some point inside the system we have a small constructional change, e. g. an increase in the curvature of the refracting surface  $\Gamma_i$  which becomes  $\Gamma_i^*$ . Let  $N_i$  and  $N_i'$  be the refractive indices of media on the object and image side of  $\Gamma_i$  respectively. The ray path  $O\Pi P_i \Pi' O'$  becomes  $O\Pi P_i^* \Pi'^* O'^*$ .  $\Pi'$  and  $\Pi'^*$  are the points of incidence of the original and varied rays at the exit pupil reference sphere whose centre is at  $o'$ , the original paraxial focus.  $P_i^* Q_i$  is the perpendicular dropped from  $P_i^*$  to the ray  $P_i \dots \Pi'$ .

\* Now at Taylor, Taylor & Hobson, Ltd, Leicester.



If by the theorems of MALUS and FERMAT the paths  $P_i^* \dots \Pi'^*$  and  $Q_i \dots \Pi'$  are equal, the change in wavefront aberration taken as the departure of the wavefront from the reference sphere centred at  $o'$  is simply —  $[N_i(PP^*) - N_i'(PQ)]$ . The reversal of sign is in accordance with the convention that a longer path along the ray, which means a wavefront retarded with respect to a fixed reference sphere, gives negative-sign aberration. Reference to  $o'^*$ , the new Gaussian focus, then requires adding the extra path  $N'\sigma$  travelled along the ray between the reference spheres centred at  $o'^*$  and  $o'$  respectively ( $N'$  is the refractive index of the image space medium).

There are two basic approximations here which break down to a greater or lesser extent according to the position in the optical system of the constructional change, its magnitude and nature, and according to the position of the exit pupil. The first approximation is the assumption of neighbouring paths already mentioned. The second assumption is that the original and varied rays at the exit pupil are so near that the change of focus term (path between reference spheres centred at  $o'$  and  $o'^*$ ) may be measured indifferently along either ray.

The first approximation becomes much less severe when a point is found in image space on the original ray which has the same optical distance from  $P_i^*$  and  $Q_i$  to an accuracy which includes the second power of the angle between the rays. Such a point is found by tracing the paraxial tangential fan about the original ray as axis, which marks the progress of the wavefront element  $P_i^* Q_i$  of aperture  $u_i'$ , now considered no longer a first order small quantity. The trace of such a paraxial fan commonly used in astigmatism calculation involves very much less calculation than the retrace of a trigonometric ray through the changed system. We are at the same time led to the elimination of the second assumption. The new focus  $O'^*$  is to the same approximation the point where the tangential fan ray intersects the axis. If we drop all reference to Gaussian foci as such, we can easily refer the aberration change, given with respect to a reference sphere centred at  $O'^*$ , to a sphere centred at  $O'$ , the original ray focus. The latter focus has been called (HOPKINS, 1952), the invariant focus, and has the advantage of giving a unique value for the wave front aberration, i. e. as the wavefront progresses along the ray, its departure from a reference sphere centred at  $O'$  remains invariant. All reference to the exit pupil of the system is thus avoided and the second approximation is circumvented.

The technique of correction is applied as follows. A sufficient number of trigonometric rays, axial and extra-axial, are traced, say 15 : marginal, sub-marginal and zonal on axis and for 2 positive and 2 negative object heights, (the latter are traced as « T-rays », (see below) relative to trigonometric pupil rays). The correction of the system is satisfactory when the Optical Path Differences, axial — ray path, (henceforth O.

P. D.) at these 15 foci are simultaneously zero. Differential changes are calculated on each of these rays and so combined as to minimize the 15 quantities. If large changes are involved which modify the basic structure of the system or give rise to completely different ray paths, it will be necessary at a certain stage to retrace the system and repeat the process. Simultaneously zero O. P. D.'s at all invariant foci would result in spherical wavefronts centred at paraxial and paraxial-tangential foci. Thus for a satisfactory result, attention must be given in the earlier stages of the design to the requirement that the system should have as far as possible a flat tangential field (as given by conventional tangential fans).

We shall develop expressions for the aberration increments working for simplicity first for the axial case, showing later how a small number of auxiliary calculations make them adaptable to the extra-axial case. First we must consider the change in O. P. D. at a surface when the ray path through it is varied.

**III. Effect on aberration at subsequent refracting surface of a constructional change.** —  $\Gamma$  is a refracting surface (pole A) between media of indices  $N, N'$ . PO, after refraction (henceforth a. r.)  $PO'$ , is a ray through the surface, P being its point of incidence. Reference spheres are drawn through P with O and  $O'$  as centres intersecting the axis at R and  $R'$  respectively.  $AR \equiv \epsilon$ ,  $AR' \equiv \epsilon'$ , with the sign convention that  $\epsilon(\epsilon')$  is negative when at the axial intersections the refracting surface is in advance of the reference sphere and vice versa. Then :

$$O.P.D.O \dots O' = N(AO) - N'(AO') - N(PO) + N'(PO') = -N(AR) + N'(AR') = \Delta(N\epsilon),$$

where  $\Delta$  denotes the difference (expression a. r.) — (expression b. r.).

We now take the ray PO, a. r.  $PO'$ , to be that traced through the unmodified optical system. A constructional change has been taken at surface  $\Gamma_i$  (fig.1) before the surface  $\Gamma$  of figure 2  $P_i^* Q_i$  and  $u_i'$  have been obtained there and a tangential fan has been traced with these quantities as starting data. The ray from this fan is (in fig. 2)  $P^* O^*$  a. r.  $P^* O'^*$ . It determines with the trigonometric ray foci  $o_i, o_i'$ ; apertures  $u_i, u_i'$  and incidence heights  $y_i, y_i'$ . We call

$$P_{oi} \equiv t, P_{oi}' \equiv t'; PO \equiv S, PO' \equiv S'; \\ < POA \equiv U, < PO'A \equiv U';$$

and the perpendiculars from  $o_i$  and  $o_i'$  to the axis respectively  $\rho_i, \rho_i'$ .

We have equality of optical paths between  $o_i$  and  $o_i'$  along the two rays and draw reference spheres through  $o_i$  and  $o_i'$  with centres O,  $O^*$  and  $O', O'^*$  respectively. The spheres determine axial intercepts  $BB^*$  and  $B'B'^*$ .

Evidently the change in O. P. D.,  $O^* \dots O'^*$  (via axis) — via  $P^*$ ) from O. P. D.  $O \dots O'$  (via axis) — (via P) is





$$1 \text{ b) } (N' \varepsilon')^* - (N' \varepsilon') = -\frac{N'}{2} \rho'_t u'_t$$

and tangential focus at infinity:  $t' \rightarrow \infty$ ,  $u'_t = 0$  as

$$(1 \text{ c}) \quad (N' \varepsilon')^* - (N' \varepsilon') = N' y'_t \frac{\sin U'}{1 + \cos U'}.$$

The expression (1 a) however does not have to be computed at every surface in the system since

$$[(N \varepsilon)^* - (N \varepsilon)]_{j+1} = [(N' \varepsilon')^* - (N' \varepsilon')]_j$$

where  $j$  and  $j + 1$  denote consecutive surfaces. In fact after the calculation has been completed at the surface whose constructional parameters are varied, say surface  $i$  we have for the complete result,

$$(2) \delta(\text{OPD}) = \delta(\text{OPD})_i + [N(S - t) \Phi(\sin U, u_t)]_{i+1} - [N'(S' - t') \Phi(\sin U', u'_t)]_k,$$

where  $k$  denotes the last surface. The difference of the last two terms is the quantity assumed vanishingly small in the approximation of FERMAT's Theorem (§§ I, II). It is the complete expression for the O. P. D. change subsequent to the modified surface.

It remains to find expressions for the parameter derivatives of the O. P. D. at the surface which is modified. The cases to be considered are : (1) a change in surface curvature and " bend " of a component ; (2) a change in position of a surface or component ; (3) a change in refractive index of a medium.

**IV. Change in curvature of refracting surface and « bend » of component.** — In figure 3 we consider the refracting surface i of figure 1 (§ II). Reference spheres are drawn as in figure 2 with  $O'$  and  $O'^*$  as centres

but intersecting in  $P^*$  instead of  $o'_1$ . The corresponding axial intersections are  $R^*$  and  $R$ . We have

$$\varepsilon^* - \varepsilon = PP^*, \quad \varepsilon'^* = AR^*, \quad \varepsilon' = AR - PQ.$$

Therefore

$$(3) \Delta(N_\varepsilon)^* - \Delta(N_\varepsilon) = N'(PQ) - N(PP^*) - N'(RR^*)$$

Q is again the foot of the perpendicular from P\* to PO', which may be confused with the portion of the reference sphere centre O' through P\*. Q\* is the intersection of reference sphere centre O'\* with PO'.

It follows from what has been said about invariant foci that  $N'(\text{RR}^*)$  is equal to the term  $N'(\text{BB}^*)$  of (1) § III, when it is remembered that

$$P^* o'_t = \begin{cases} Q^* o'_t \\ Q o'_t \end{cases} + O(u'_t)^2.$$

Thus the last term of (3) cancels with the second term of (2) § III. We proceed to derive an expression for the first two terms of (3) and for the  $y'_i$  and  $u'_i$  which mark the tangential fan to be traced through the remainder of the system.

In what follows rigid adherence to the same degree of approximation has been abandoned. Thus terms depending on the square of the parameter change are neglected where these can be considered random (1).

There is however one case where squared terms are of particular importance i. e. in the "bend" of a component near minimum spherical aberration, the vertex of the parabola (RUMSEY, 1951). The non-random square terms are then separated out and independently computed.

(<sup>1</sup>) Apart from simplification of formulae there is an advantage in the results being directly scalable.

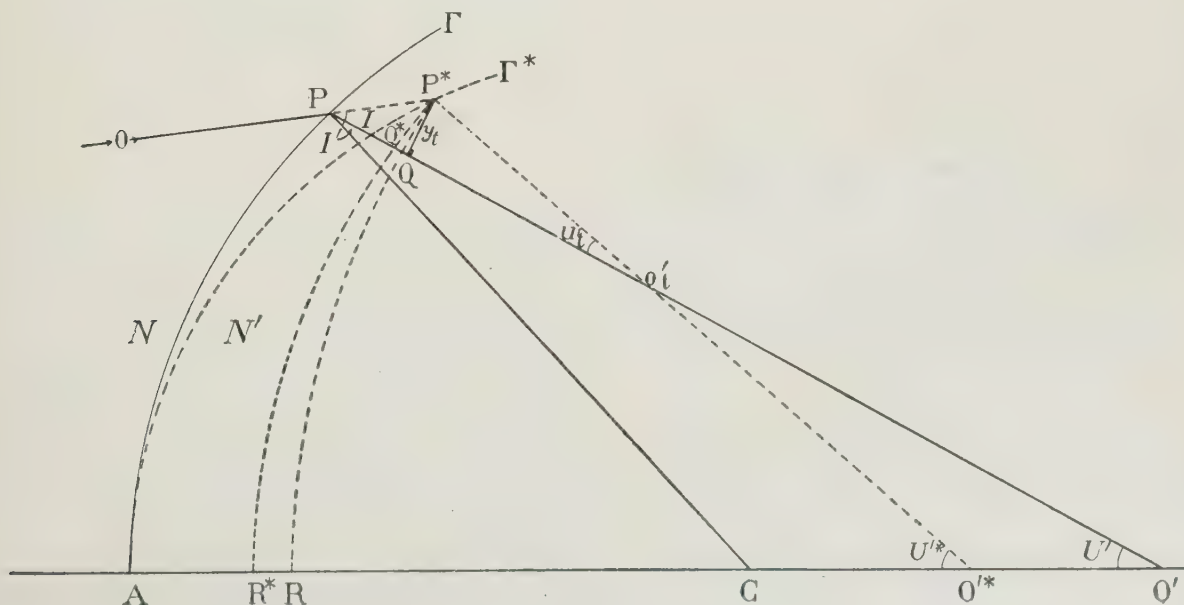


FIG. 3.

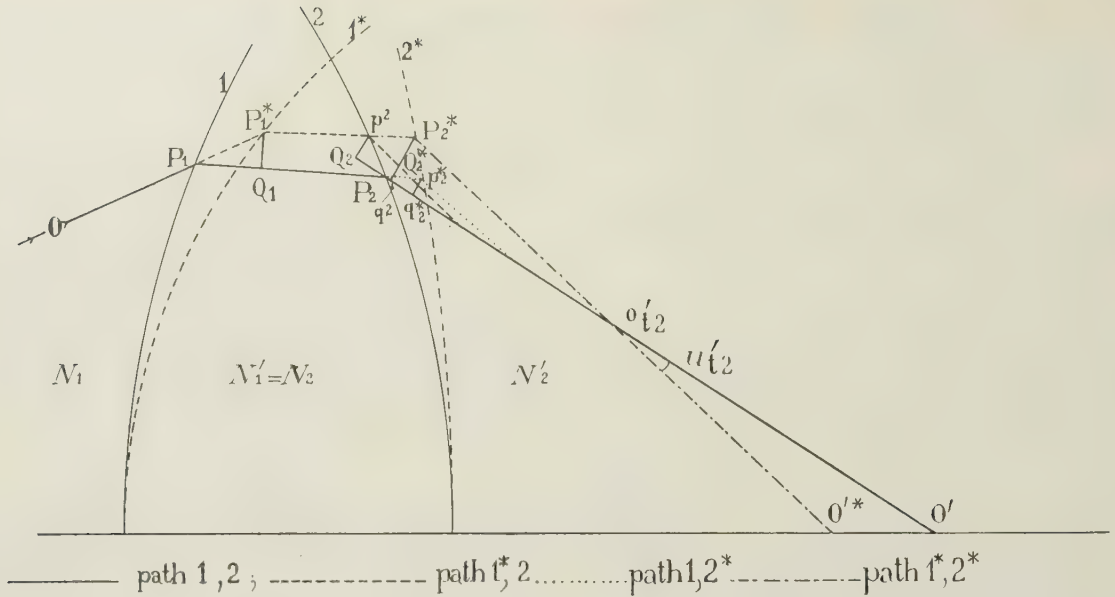


FIG. 4 a.

Using the formulae of the ray trace scheme given below (Appendix II) we have :

$$\varepsilon = \frac{2h \sin I}{G}, \quad G \equiv (1 + \cos U)(1 + \cos I) + \sin U \sin I$$

which gives upon differentiation with respect to  $c$ , after rearrangement,

$$(4) \quad PP^* = \varepsilon^* - \varepsilon = \frac{2h^2 \delta c}{G \cos I} \cdot \frac{(1 + \cos U)(1 + \cos I)}{G}$$

but  $PQ = PP^* \cos(I - I')$ , and with the aid of  $\Delta(N \sin I) = 0$

$$(5) \quad N'(PQ) - N(PP^*) = (\varepsilon^* - \varepsilon) \cos I \Delta(N \cos I),$$

while

$$(6) \quad y'_t = PP^* \sin(I - I') = (\varepsilon^* - \varepsilon) \frac{\sin I}{N'} \Delta(N \cos I).$$

Let  $I^*$  (b. r.),  $I'^*$  (a. r.) be the angles of incidence of  $OP^*$  a. r.  $P^*O'^*$  with  $\Gamma^*$ . As  $U$  is unchanged,  $u'_t$  is given by the change in the deviation of the incoming ray. Therefore  $u'_t = -\Delta(I)^* + \Delta(I) = -\delta\{\Delta(I)\}$ .

But in the form of ray trace given in Appendix II

$$\sin I = hc - \sin U$$

where  $h$  is the perpendicular from  $A$  to  $OP$ .

Therefore

$$(7) \quad \delta I = \frac{h \delta c}{\cos I} = -\Omega$$

while because of  $\Delta(N \sin I) = 0$ ,  $\delta I' = \frac{N \cos I}{N' \cos I'} \delta I$ .

Therefore

$$(8) \quad u'_t = \Omega \left( \frac{N \cos I}{N' \cos I'} - 1 \right)$$

(8) may be interpreted as follows: The change  $\Gamma, \Gamma^*$

means at the point of incidence of the ray  $OP$  a change in the tilt of refracting surface relative to incoming wavefront. If we tilt instead the wavefront by  $\Omega = -\delta I$ ; the new ray is treated by  $\Gamma$  as was  $OP$  by  $\Gamma^*$ . Thus a fan with aperture  $\omega = \Omega$ ,  $y_t = 0$  is traced through the surface and the resulting  $\omega'$  is referred to the ray  $PO'$  by subtracting from it the tilt angle  $\Omega$ . If the subsequent surfaces are unchanged we use  $u'_t$  and  $y'_t$  of (6) to trace a fan through to the end of the system <sup>(2)</sup>.

This interpretation of (8) leads directly to a method of calculating the "bend" of the component. Such a "bend", we take not as equal increments  $\delta c$  imposed on all surfaces but as a constant tilt  $\Omega$  of the incoming ray. The effective changes of curvature are then given for each surface by

$$(7a) \quad \delta c = -\frac{\Omega \cos I}{h}.$$

Such a procedure is known from first order calculations based on a paraxial ray. However, while in the latter case a constant increment  $y\delta c$  ( $y$  = paraxial incidence height) gives exactly unchanged deviation of the paraxial ray this is only approximately so in our case. Further, while in the paraxial case the increment  $y\delta c$  may be scaled for other rays in the finite case  $\Omega$  can be constant for only one ray and has to be obtained anew from (7) for the other rays.

In figure 4a we consider two consecutive surfaces of a component. They are "bent" by  $\Omega_1$ , and  $\Omega_2$  corresponding to  $\delta c_1$  and  $\delta c_2$  [by (7)]. The  $\delta c_1$  and  $\delta c_2$

<sup>(2)</sup> In working out  $y_{t+1} = y'_t - Du'_t$  we are justified in the use of  $D$ , the distance along  $PO'$  to the next surface in place of  $D - PQ$  since only a square term is thereby neglected.



are in turn derived from a constant  $\Omega$  [by (7a)] at some other ray. The path 1, 2 through the unbent component becomes 1\*, 2\* through the bent component. We require the difference in path from  $P_1$ , via 1, 2 and via 1\*, 2\* to a point  $o'_t$  which is found from a "tangential fan" traced in the following manner.

a) A tangential fan with  $\omega_1 = \Omega_1$ ;  $y_{t_1} = 0$  is traced through 1.

b)  $u'_{t_1} = \omega'_1 - \Omega_1$ .

c)  $y'_{t_1}$  is obtained from (6).

d)  $y_{t_2} = y'_{t_1} - D_{1,2} u'_{t_1}$  where  $D_{1,2} = P_1 P_2$ .

e)  $\omega_2 = u'_{t_1} + \Omega_2$ .

f) A tangential fan with  $\omega_2, y_{t_2}$  is traced through 2.

g)  $u'_{t_2} = \omega_2 - \Omega_2$ .

h)  $y'_{t_2}$  is found by adding to the value  $y_{t_2} \frac{\cos I'_2}{\cos I_2}$  derived from the tangential fan in (f) the result of applying equation (6) at surface 2.

i) If the component subjected to "bending" ends at surface 2, we have the required point  $o'_t$  (in this case  $o'_{t_2}$ ) at  $y'_{t_2}$  from surface 2, where  $y'_{t_2} = t'_2 u'_{t_2}$  is the value obtained in (h).  $t'_2, u'_{t_2}$  are then directly used in equation (2) of § III. If "unbent" components or surfaces follow, a conventional fan is traced on to the end of the system and (2) of § III is applied in image space.

j) If more surfaces remain in the component subjected to "bending" or if the following component is also "bent" the procedure (e), (f), (g) and (h), for such surfaces is repeated.

Referring back to figure 4a we divide path difference (N. B. sign convention in accordance with § II). 1,2 — 1\*, 2\* into components 1, 2 — 1\*, 2 and 1\*, 2 — 1\*, 2\*, that is, we consider the change in curvature of each surface in turn. The procedure outlined above in steps (a) to (j) is to relate to this construction.

We have

$$1,2 - 1^*,2 = [P_1 - Q_2] - [P - P_2] = N'_1(P_1 Q_1) - N_1(P_1 P_2^*)$$

and

$$1^*,2 - 1^*,2^* = [P_2 - Q_2^*] - [P_2 - P_2^*] = N'_2(P_2 Q_2^*) - N_2(P_2 P_2^*) + O(\delta c^2)$$

but

$$[P_2 - Q_2^*] = [Q_2 - q_2].$$

Hence

$$1,2 - 1^*,2^* = [P_1 - q_2] - [P_1 - P_2^*] = \sum_{1,2} (\varepsilon^* - \varepsilon) \cos I \Delta(N \cos I).$$

The point  $o'_{t_2}$  is equidistant from  $P_2^*$  and  $q_2$  and  $P_2^* q_2$  is the  $y'_{t_2}$  obtained in step (h), since

$$P^* q_2 = p_2 Q_2 + p_2^* q_2^* + O(\delta c^2).$$

Also  $\angle P_2^* o'_{t_2} q_2$  is the required angle  $u'_{t_2}$  except for a term depending on  $\delta c^2$ , being the addition of the angles made by path 1\*, 2 with 1, 2 and 1\*, 2\* with 1\*, 2. Here however step (f) does not correspond exactly to the construction of figure 4a viz. the fan marking path 1\*, 2 relative to path 1, 2 through surface 2 should be traced with starting data  $y_{t_2}, u_{t_2}$  not  $y'_{t_2}, \omega$  as in (f). The additional path difference involved can be seen from figure 4b to be a quantity dependant on  $\delta c^2$ . OPO', figure 4b, is a trigonometric ray through a surface.  $P^* o_t$ , a. r.  $P^* o'_t$  is a fan with apertures  $\omega_t$ , a. r.  $u'_t$ .  $o_t^* P^*$ , a. r.  $o'_t P^*$  is a fan with aperture  $\omega$ , a. r.  $\omega'$ .  $\omega = u_t - \Omega$ ,  $\omega' = u'_t - \Omega$ . We require the change in the path of  $o_t^*$  to  $o'_t^*$  from that of  $o_t$  to  $o'_t$  via P — via P\* in each case. The consideration is analogous to that of § III and leads to the form (1b) § III with  $y_t$  taking the place of  $\rho_t$  and  $\Omega$  the place of  $u_t$ .

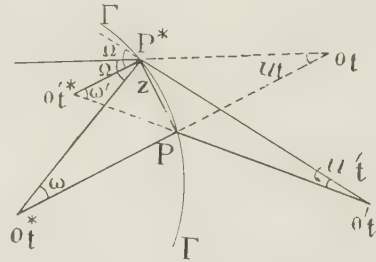


FIG. 4 b.

The term becomes

$$-\frac{1}{2} \Omega \Delta(N y_t) = -\frac{1}{2} \Omega z \Delta(N \cos I),$$

where  $z = \frac{y_t}{\cos I} = \frac{y'_t}{\cos I'}$ , and has to be subtracted

from the path difference 1,2 — 1\*, 2\*. Importance is attached to it as a non-random square term. One of the characteristics of a "bend" is the progressive rise or drop in incidence heights of a ray through the component according to the sign of  $\delta c$ . Thus  $\Omega z \Delta(N \cos I)$  will never average out. It becomes in fact the controlling term at the minimum of the spherical aberration parabola where  $\partial(O. P. D.)/\partial \Omega = 0$ . Summing up, we have for a "bend" of a component, when the calculations are carried through as above,

$$(9) \quad \Delta(N \varepsilon^*) - \Delta(N \varepsilon) = \sum_i (\varepsilon^* - \varepsilon) \cos I \Delta(N \cos I) + \sum_i \frac{1}{2} z \Omega \Delta(N \cos I) - [N'(S' - t') \Phi(\sin U', u'_t)]_k$$

V. Displaced component or surface. — The problem of an axial displacement of a component or surface is very simply treated by considering a con-





$\delta u'_{12a} = \frac{N'_2 \delta N}{N_2 \cos I_2} (u'_{12a} + z_2 c_2)$  may be added if significant. The difference in paths 1, 2 and 1\*, 2\* (respecting sign convention, viz.  $\delta N > 0$  means a longer ray path) then becomes

$$\begin{aligned} \Delta(N_1 \epsilon_1)^* - \Delta(N_1 \epsilon_1) + \Delta(N_2 \epsilon_2)^* - \Delta(N_2 \epsilon_2) = \\ = \delta N(\epsilon'_1 - \epsilon_2) + N'_1 \delta \epsilon'_1 + N'_2 \delta \epsilon'_2 - N_2 \delta \epsilon_2 \end{aligned}$$

where the  $\delta \epsilon$  values are the axial intercepts marked in figure 6. But by the property of invariant foci ( $O'_1, O_1^*$ )  $\delta \epsilon'_1 = \delta \epsilon_2$ , while  $N'_2 \delta \epsilon'_2$  again cancels with the second term of (2) § III and  $\epsilon'_1 - \epsilon_2$  is equal to d-D, i. e. axial separation of surfaces minus separation along the ray. The complete result reduces to

$$\begin{aligned} (14) \quad \Delta(N\epsilon)^* - \Delta(N\epsilon) = \\ = (d - D) \delta N - [N'(S' - t') \Phi(\sin U', u'_t)]_k \end{aligned}$$

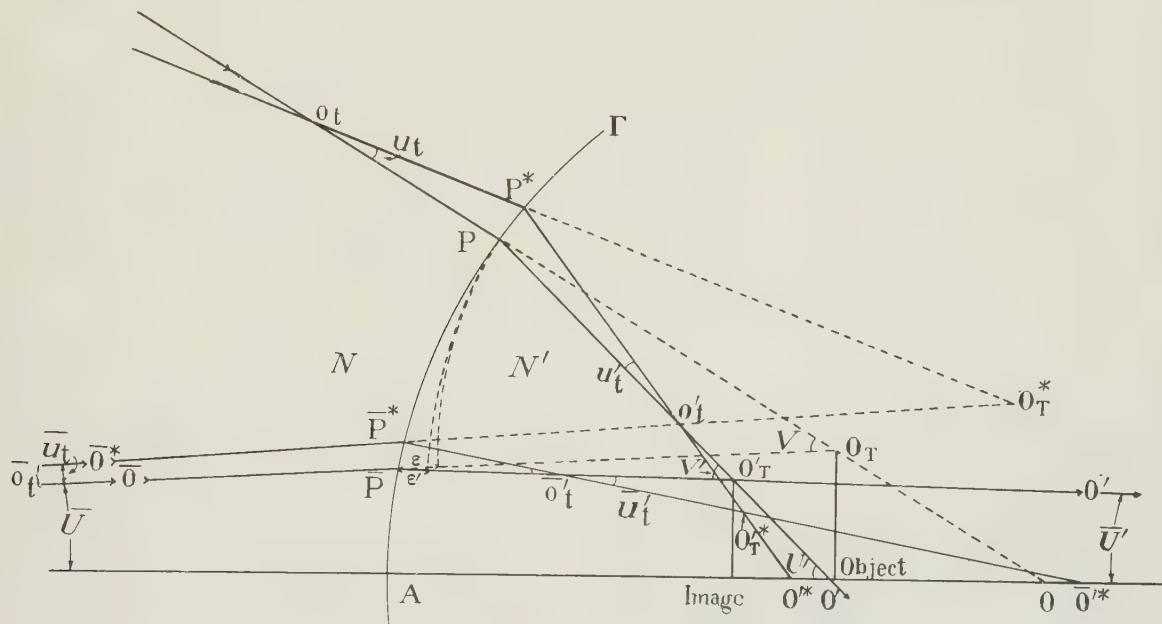


FIG. 7.

**VII. The extra-axial meridional case.** — It is conventional practice to treat the extra-axial (meridional) aberrations by tracing rays through the point of the object field under consideration which are in effect axial rays imaging fictitious object points. Formulae for the extra-axial O. P. D. have been given by CONRADY (see HOPKINS, 1950) which introduce data from such rays together with data from the pupil ray (which images the entrance pupil). As in off-axis imagery one is only concerned with the relative paths along the two rays, it is reasonable that the extra-axial ray should be traced in the first place relative to the pupil ray (T-ray) and not the axis. (Formulae for such a ray trace scheme are given in Appendix II). If such a trace includes quantities  $\epsilon, \epsilon'$  defined as before, § 2, but with the axial intercepts becoming intercepts

on the pupil ray, the O. P. D. is again given by  $\Delta(N\epsilon)$ .

In figure 7 we show a typical pupil ray, T-ray through a surface and the variation of both rays. The constructions and notations are obvious extensions of the axial case.

If we now repeated the considerations of § 3 we should arrive at an expression analogous to  $N(S - t) \Phi(\sin U_1 u_1)$  in the form

$$\begin{aligned} (15) \quad N[(S - t) \Phi(\sin V, u_t) + \\ + (L - \bar{t}) \Phi(\sin V, \bar{u}_t) - \frac{\bar{u}_t}{1 + \cos V} \\ \{ 2[(S - t) - (L - \bar{t})] u_t + (L - \bar{t}) \bar{u}_t (1 - \cos V) \}] \end{aligned}$$

The proof of (15) is laborious and is not given here. Instead both the formulae are treated again in Appendix I by a more elegant method. So far as the considera-

tions of §§ IV, V, VI, are concerned the simplest procedure is to extract from the traced T-rays, the quantities used in these sections, which are then considered relative to the axis as before. Thus after the obvious derivation of  $\sin U, \cos U$  from  $U = U + V$  we use  $h = (\sin U + \sin I)/c$ , all other quantities appearing explicitly in the T-trace. The pupil ray itself is an axial ray and its variation requires no additional considerations.

**§ VIII. Application and discussion of numerical results.** — The system of differential correction has been applied to two designs of microscope objectives: A 2 m. m. Oil Immersion Achromat of 1.30 N. A., and a 4 mm dry Achromat of 0.85 N. A.

In interpreting Tables I and II we should note the

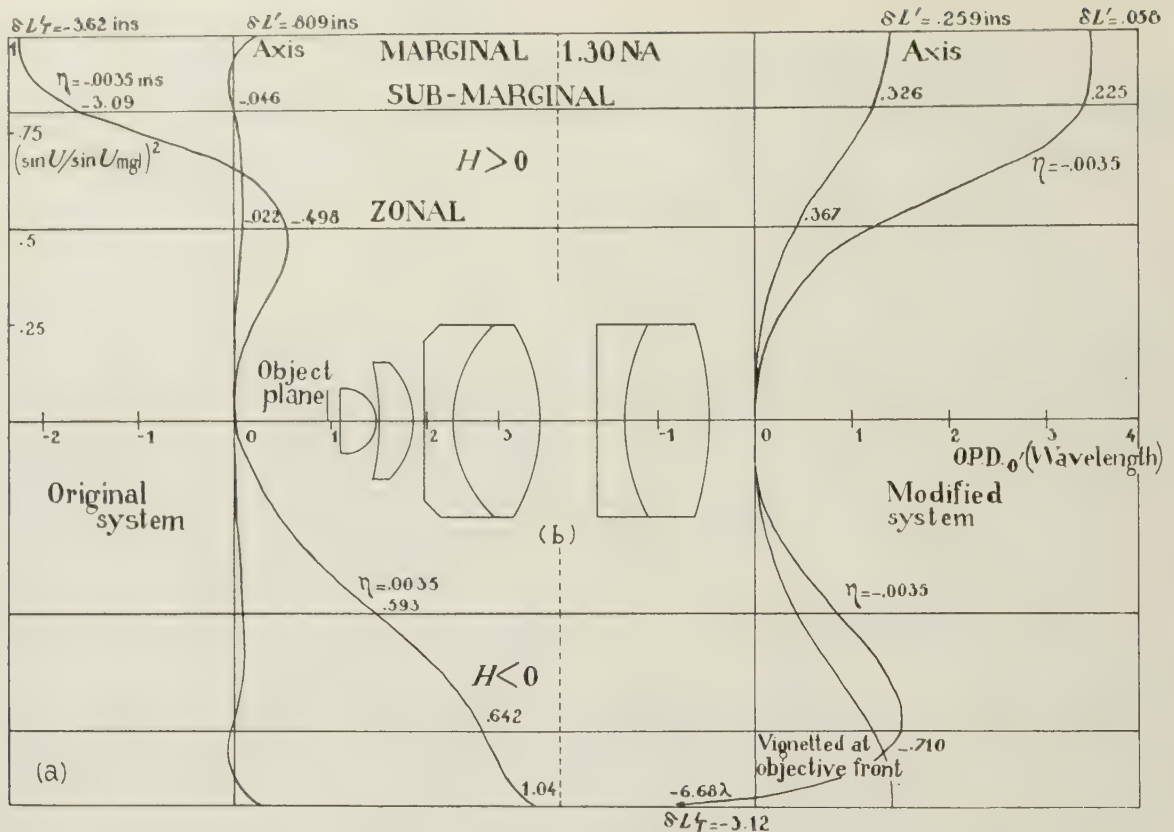


FIG. 8.

following: The O. P. D.'s at invariant foci are always opposite in sign to those at the paraxial focus. The reason for this is that the change of focus term predominates over the path difference along the ray. Thus an undercorrection (positive at the paraxial focus) is indicated by a negative O. P. D. at the invariant focus and vice-versa. Comparing at a given aperture, the O. P. D. on axis with those derived from the T-rays, we see that the values for negative and positive fields (denoted by the sign of  $H$ , the LAGRANGE invariant) lie on either side of the axial value. This indicates an odd term in the extra-axial O. P. D. which expresses the coma. On the other hand the non-paraxial astigmatism and excess of spherical aberration on the principal ray over that on the axis are expressed by the mean of the extra-axial O. P. D.'s. (at invariant foci) being greater (overcorrection) or smaller (undercorrection) than the corresponding axial value.

In Table I results are quoted for seven constructional changes of the first objective. They are taken in two stages and so chosen as to leave the convergence of the marginal ray unchanged (i. e.  $\Sigma u'_{t_k} = 0$ ). A comparison is made by retracing the system after the first three changes. The remaining four changes are calculated on the intermediate state and compared with a retrace of the objective in its final form.

In its original state (see also fig. 8 a) the objective shows two defects: 1) the axial wavefront, while very close to the paraxial reference sphere (in fig. 8 a the  $y$ -axis), deviates rapidly at the margin viz: has a rapid change in gradient; 2) the coma is overcorrected or outward (negative in our convention i. e. wavefront for  $H > 0$  retarded with respect to reference sphere while in advance for  $H < 0$ ). In Table I, 1) is seen from the large negative O. P. D. at the marginal invariant focus; 2) from the fact that the O. P. D. at the sub-marginal T-ray focus is more positive for  $H > 0$  than for  $H < 0$ .

The result of the modification has been to eliminate the marginal under-correction at the sacrifice of a small amount of zonal undercorrection. The latter is largely taken care of by the choice of best focus. Thus in figure 8 a the wavefronts (referred to paraxial foci) before and after modification are plotted against the square of the aperture. The paraxial reference sphere (also the reference sphere for all foci lying in the Gaussian image plane) is taken as coincident with the  $y$ -axis. The reference sphere at any other focus is a straight line inclined to it. Hence the asphericity of the wavefront is judged by its approximation to a straight line.

The coma has been largely removed. The O. P. D.'s



TABLE I

Objective : 2 mm. 1.30 N. A.							
Achromatic Lens.	Axis			Fields			
	Mgl.	S. Mgl.	Zonal	$H > 0$		$H < 0$	
				Mgl.	S. Mgl.	Mgl.	S. Mgl.
Original System O. P. D.	— 00 0094	00 0004	00 0003	00 0191	00 0144	— 00 0051	00 0003
$\delta N'_1 = \delta N_2 - .02$ .....	0577	0216	0040	0949	0426	0337	0101
$\delta c_5$ — .5117 .....	— 0420	— 0202	— 0047	— 0692	— 0320	— 0226	— 0114
$\delta c_{11}$ .0356 .....	0005	0003	0001	0003	0002	0024	0021
$\delta(OPD)$ by differential Correction .....	0162	0017	— 0006	0260	0108	0135	0008
Intermediate System							
OPD* (by differential Correction) ....	0068	0021	— 0003	0451	0252	0084	0011
OPD* (retraced) .....	0007	0009	— 0001	0191	0128	0168	0064
$\delta N'_2 = \delta N_3 - .01127$ .....	0314	0119	0023	0476	0184	0233	0078
$\delta c_5$ — .2883 .....	— 0311	— 0132	— 0030	— 0527	— 0212	— 0177	— 0076
$\delta z_{8-9}$ — .02 .....	— 0003	— 0003	— 0003	— 0079	— 0078	0	— 0009
$\delta c_{11}$ .079618 .....	0012	0009	0003	— 0022	— 0018	0032	0026
$\delta(OPD)$ by differential Correction....	0012	— 0007	— 0007	— 0152	— 0124	0088	0019
Modified System							
OPD* (by differential Correction) ....	0019	0002	— 0008	0039	0004	0256	0083
OPD* (retraced) .....	0002	— 0003	— 0012	0073	0057	0233	0089

TABLE II

Objective : 4 mm. 85 N. A.							
Achromatic Lens.	Axis			Fields			
	Mgl.	S. Mgl.	Zonal	$H > 0$		$H < 0$	
				Mgl.	S. Mgl.	Mgl.	S. Mgl.
Original System O. P. D.	00 0128	00 0066	— 00 0048	00 1026	00 0346	— 00 0536	— 00 0177
$\delta z_{0-1}$ — .00612 .....	2354	0940	0124	2562	0956	2169	0956
$\delta z_{2-3}$ .005942 .....	— 0471	— 0173	— 0016	— 0544	— 0181	— 0431	— 0180
$\delta c_3$ — .45 .....	— 0562	— 0212	— 0019	— 0751	— 0277	— 0415	— 0117
$\delta z_{3-4}$ — .002 .....	— 0159	— 0072	— 0008	— 0324	— 0151	— 0059	0001
$\delta c_4$ 1.35 .....	— 1655	— 0722	— 0090	— 2181	— 0798	— 1046	— 0628
$\delta z_{4-5}$ — .013965 .....	— 0550	— 0241	— 0024	— 1113	— 0516	— 0180	— 0029
$\delta N'_5 = \delta N_6 - .03095$ .....	1583	0762	0111	2168	0930	1097	0620
$\delta c_6$ — .9375 .....	— 1085	— 0522	— 0082	— 1605	— 0608	— 0865	— 0513
$\delta c_7$ — .16 .....	— 0200	— 0120	— 0026	— 0261	— 0146	— 0137	— 0101
$\delta c_9$ — .05 .....	— 0034	— 0024	— 0007	— 0016	— 0010	— 0046	— 0036
$\delta(OPD)$ by differential Correction .....	— 0379	— 0144	0015	— 1543	— 0509	0361	0175
Intermediate system.							
OPD* (by differential Correction) .....	— 0251	— 0078	— 0033	— 0517	— 0163	— 0175	— 0002
OPD* (retraced) .....	— 0001	0011	— 0028	0343	0155	— 0323	— 0126
$\delta z_{0-1}$ — .0011 .....	0320	0150	0024	0324	0158	0324	0155
$\delta z_{2-3}$ .0084 .....	— 0383	— 0188	— 0028	— 0393	— 0212	— 0395	— 0201
$\delta z_{4-5}$ — .0025 .....	— 0051	— 0024	— 0002	— 0112	— 0066	— 0003	0010
$\delta c_8$ — .3 .....	— 0212	— 0114	— 0020	— 0328	— 0159	— 0133	— 0082
$\delta N'_8 = \delta N_9 - .0199$ .....	0394	0242	0053	0502	0276	0324	0218
$\delta(OPD)$ by differential Correction....	0068	0066	0027	— 0007	— 0003	0117	0100
Modified System.							
OPD* (by differential Correction) .....	0067	0077	— 0001	0336	0152	— 0206	— 0026
OPD* (retraced) .....	0020	0050	— 0006	0347	0195	— 0309	— 0099

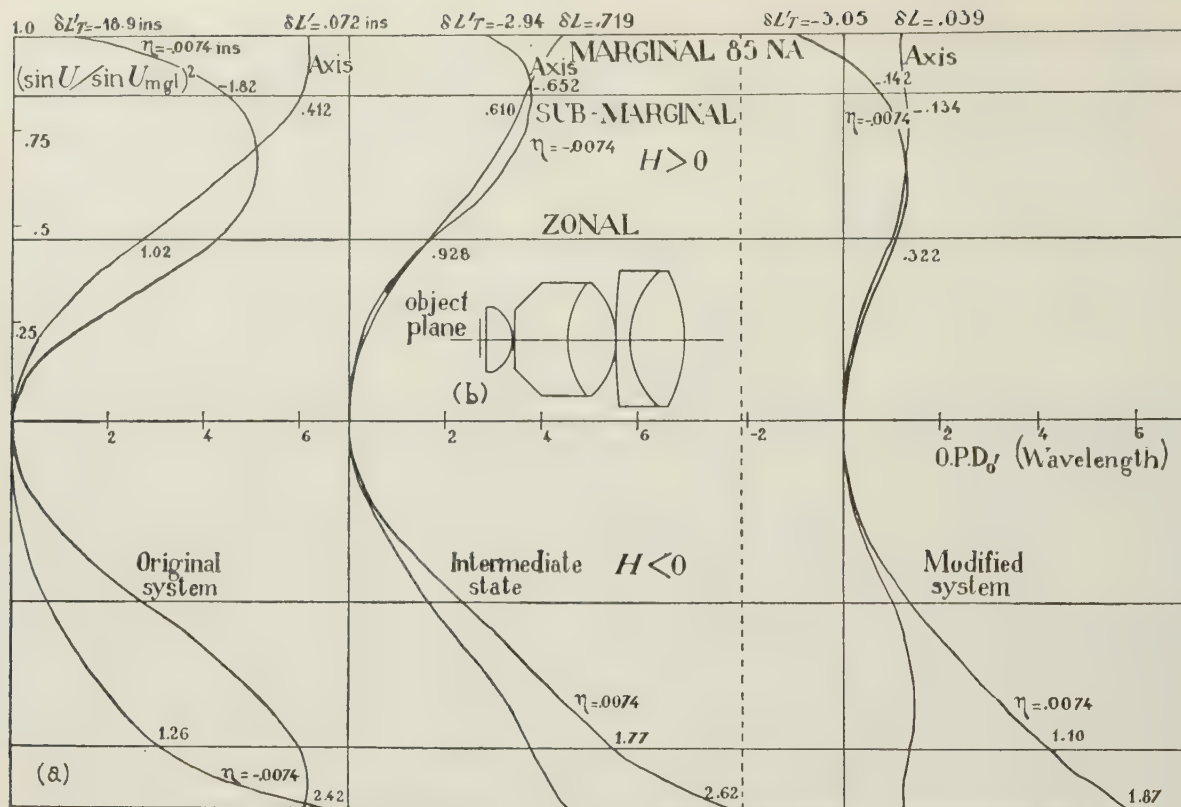


FIG. 9.

at the submarginal T-ray foci are nearly equal for  $H > 0$  and  $H < 0$ . For the marginal T-rays there is now a comatic under-correction, viz. O. P. D. more positive for the  $H < 0$  than  $H > 0$ . There is thus a zonal variation of coma which precludes its complete elimination. To the tangential focus (which includes both PETZVAL curvature and astigmatism dependant on the square of the aperture) again corresponds a straight line as reference sphere. This will have an inclination with respect to the  $y$ -axis dependant on the field curvature (obviously positive) of the objective. The asphericity of the extra-axial wavefronts is judged by their departure from such a line. Thus we note (fig. 8 a) as well as the residual coma, an overcorrection dependant on even powers of the aperture. This however is tolerated since it assists the flattening of the effective field of the objective, while the worst of the comatic under correction at the margin, is eliminated by vignetting at the objective front.

$\delta L'(\delta L'_r)$  values (all measured from the Gaussian image plane) are quoted for comparison with the conventional method of characterising the correction of an objective.

In Table II (fig. 9) the technique of differential correction is applied to a 4 mm. 85 N. A. dry achromatic objective. Such an objective is generally considered over-apertured and presents considerable difficulties

to the designer (see CONRADY unpublished notes on microscope objective design). The objective chosen as the original system was severely under corrected at the zone, being corrected to a best focus considerably shorter than the paraxial focus. However, except for some over corrected coma, its field correction, notably field curvature, was not excessive. The main aim of the differential changes was thus to reduce the zonal under correction while maintaining as large a working clearance as possible. To be able to do this, the first requirement was the removal of the marginal over correction on axis. This was achieved in the first ten differential changes leading to the intermediate state (fig. 9 a). A large number of compensating changes had to be taken to maintain the convergence of the axial ray. Once a state of balance was reached at the edge of the aperture, leaving it in fact slightly undercorrected, it was possible to over correct further over the whole aperture and so remove to a large extent the zonal undercorrection. The last five changes, calculated on the intermediate state, give the required result. The coma in the final form remains overcorrected but at only one third of its original value.

**Acknowledgments.** — Acknowledgment is made to Dr H. H. HOPKINS as stated in § I and unpublished



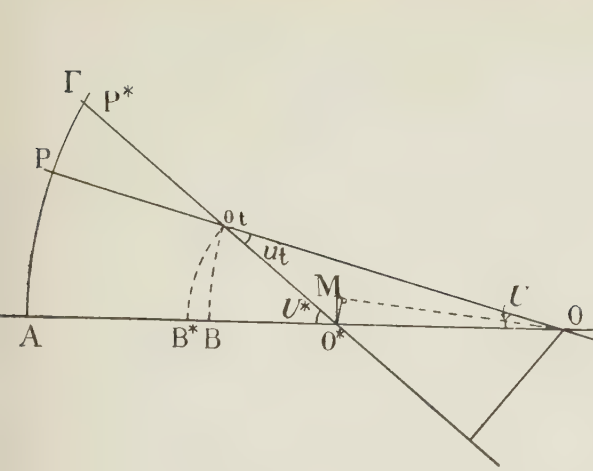


FIG. 10 a.

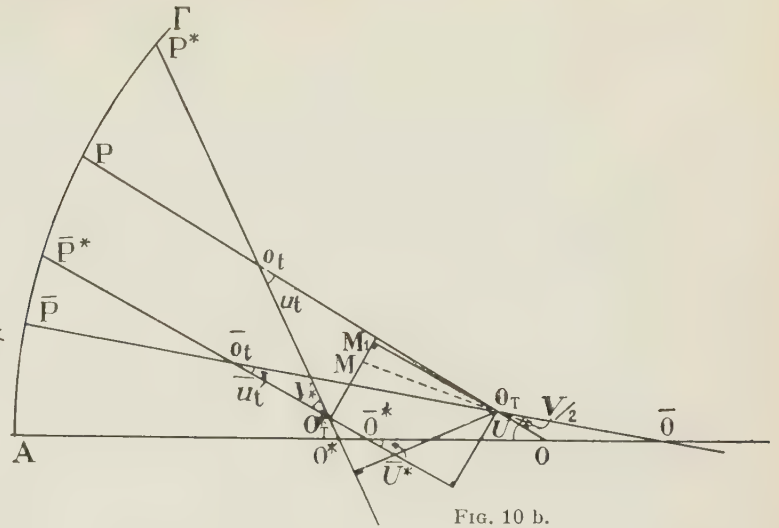


FIG. 10 b.

lecture notes of his have been consulted. The ray trace scheme of Appendix II was first developed by him for the axial case.

Acknowledgment is also made to Miss WELLINGS for assistance in computation and to the firm of COOKE, TROUGHTON & SIMMS Ltd for permission to publish.

**Appendix I.** — An alternative method of deriving the focal shift terms used in the paper is given by the following construction. In figure 10a we consider the image (or object) space of figure 2, § III, alone and draw the bisector of the angle  $U$ . The perpendicular to the axis at  $O^*$  intersects it at  $M$ .  $O^*M \equiv \sigma$ . The aberrations of an optical system are invariant not only at the ray focus, but also at any point on the bisector of the convergence angle (HOPKINS, 1952). Hence the transverse focal shift (normal to the axis),  $O^*M$ , is equivalent to the longitudinal shift (along the axis)  $O^*$  to  $O$ . From the geometry of the figure

$$\sigma = \frac{(S-t) u_t}{\sin U^*} \tan \frac{U}{2} = (S-t) \frac{u_t}{1 + \cos U} \cdot \frac{\sin U}{\sin U^*}.$$

Substituting this in the transverse focal shift formula (BLASCHKE, 1954) we have

$$-N(S-t) \frac{\sin U}{1 + \cos U} u_t \left( 1 + \frac{1}{2} \frac{\sin U}{1 + \cos U} u_t \right) \equiv \\ \equiv -N(S-t) \Phi(\sin U, u_t).$$

This is the expression we have derived for the path  $N(B*B)$  of figure 2, § III. But in § IV, figure 3 the corresponding intercept is  $R^*R$ , with reference spheres intersecting at  $P^*$  instead of  $O_t'$ . This introduces a factor into the squared term to give

$$-N(S-t) \frac{\sin U}{1 + \cos U} u_t \left( 1 + \frac{1}{2} \frac{\sin U}{1 + \cos U} u_t \cdot \frac{S-t}{S} \right).$$

Nevertheless the error committed in § IV by setting  $R^*R = B*B$  is quite negligible.

We extend the treatment to the extra-axial case.

In figure 10b  $\bar{P}\bar{O}$ ,  $PO_T$ ;  $\bar{P}^*O^*$ ,  $P^*O_T^*$ , are respec-

tively pupil and T-rays through the original and changed systems.

$$PO_T \equiv S, PO_T \equiv L, PO_t \equiv t, \bar{P}O_t \equiv \bar{t};$$

$$< P^*O_t P \equiv u_t, < \bar{P}^*O_t \bar{P} \equiv \bar{u}_t;$$

the remaining notation being as before. We require the length of the normal to  $\bar{P}^*O^*$  at  $O^*$  taken to its intersection  $M$  with the bisector of the angle  $PO_T\bar{P}(V)$ . If  $O_T M_1$  is drawn parallel to  $\bar{P}^*O^*$  then

$$\sigma \equiv OM = (L - \bar{t}) \bar{u}_t - M_1 O_T + \tan \left\{ U^* - \left( U - \frac{V}{2} \right) \right\}.$$

Substituting

$$M_1 O_T = (L - \bar{t}) \bar{u}_t \tan V^* + \left[ (S - t) u_t - \frac{(L - \bar{t}) \bar{u}_t}{\cos V^*} \right] \sin V^*,$$

and noting that

$$U^* - \left( U - \frac{V}{2} \right) = \bar{u}_t - \frac{V}{2},$$

we have

$$\sigma \sin V^* = (L - \bar{t}) \bar{u}_t \left[ \sin V^* + \cos V^* \tan \left( \bar{u}_t - \frac{V}{2} \right) \right] - \\ - (S - t) u_t \tan \left( \bar{u}_t - \frac{V}{2} \right).$$

With the same approximation as before

$$\tan \left( \bar{u}_t - \frac{V}{2} \right) = - \frac{\sin V}{1 + \cos V} + \frac{2 \bar{u}_t}{1 + \cos V}.$$

Hence

$$\sigma \sin V^* = (L - \bar{t}) \bar{u}_t \left( \frac{\sin V}{1 + \cos V} + 2 u_t - \frac{2 \bar{u}_t}{1 + \cos V} \right) + \\ + (S - t) u_t \left( \frac{\sin V}{1 + \cos V} - \frac{2 \bar{u}_t}{1 + \cos V} \right).$$

The focal shift term is given as in the axial case by substituting  $\sigma \sin V^*$  in the transverse focal shift formula. Collecting terms and making some further approximations consistent with those made in the





pupil ray. —  $\varepsilon$  is then given by the focal shift  $O_r$  to  $O_c$  along the pupil ray. (The negative sign is in accordance with the convention § III).

In deriving an expression for  $\varepsilon$  no approximation is justified and we must return to (1) of § III. The place of  $u_t$  is taken by

$$\begin{aligned} < O_c P O_r = < O_c P C + I = \\ = 2 \left( \frac{\alpha + \bar{\alpha}}{2} - U \right) - (\alpha - \bar{\alpha} + \bar{I}) + I = I + \bar{I} \end{aligned}$$

$$\text{when } -\varepsilon = S \left( 1 - \frac{\sin V + \sin(I + \bar{I})}{\sin(V + I + \bar{I})} \right).$$

Replacing  $S$  by  $\frac{h_r}{\sin V} - \varepsilon$  we have

$$\begin{aligned} \varepsilon = - \frac{h_r}{\sin V} \left( \frac{\sin(V + I + \bar{I})}{\sin V + \sin(I + \bar{I})} - 1 \right) \\ = \frac{2 h_r \sin \frac{(I + \bar{I})}{2}}{(1 + \cos V) \cos \left( \frac{I + \bar{I}}{2} \right) + \sin V \sin \frac{(I + \bar{I})}{2}}. \end{aligned}$$

For convenience of computation we multiply numerator and denominator by  $2 \cos \frac{(I - \bar{I})}{2}$  and arrive at the expression

$$(iii) \quad \varepsilon = \frac{2 h_r (\sin I + \sin \bar{I})}{G}, \quad G = (1 + \cos V)(\cos I + \cos \bar{I}) + \sin V (\sin I + \sin \bar{I})$$

Before considering the relations for the refraction of the ray we derive an alternative expression for  $Y_r$ . It is used as a check on computations. Referring to figure 11 b,

$$\begin{aligned} Y_r &= [Y - \bar{Y} + r(\cos \bar{\alpha} - \cos \alpha) \tan \bar{U}] \cos \bar{u} \\ &= r \{ \sin(\alpha - \bar{U}) - \sin(\bar{\alpha} - \bar{U}) \} \\ &= r \{ \sin(I + V) - \sin \bar{I} \} \end{aligned}$$

Or in form convenient for computations.

$$(iv) \quad E = r(\sin I \cos V + \cos I \sin V), \quad E \equiv Y_r + r \sin \bar{I}.$$

Two changes have to be considered in crossing the refracting surface  $Y_r$  to  $Y'_r$  and  $\sin V$  to  $\sin V'$ . (In the axial case only the latter). Referring to figure 11 b)

$$\begin{aligned} Y_r &= [P\bar{g} - \bar{A}\bar{g} \tan \bar{I}] \cos \bar{I} \\ &= r[\cos \bar{I} \sin(\alpha - \bar{\alpha}) - (1 - \cos(\alpha - \bar{\alpha})) \sin \bar{I}]. \end{aligned}$$

Whence

$$\begin{aligned} N[Y_r - r \cos \bar{I} \sin(\alpha - \bar{\alpha})] &= \\ &= N'[Y'_r - r \cos \bar{I}' \sin(\alpha - \bar{\alpha})] = \\ &= N \sin \bar{I} (\cos(\alpha - \bar{\alpha}) - 1), \end{aligned}$$

the righthand side being invariant (invariant in the conventional sense of a quantity which is the same a. r. and b. r.) Replacing in this  $\alpha - \bar{\alpha}$  by  $V + (I - \bar{I})$  we have

$$(v) \quad \begin{aligned} &N' Y_r = N Y_r + \\ &+ r \{ \sin V \cos(I - \bar{I}) + \cos V \sin(I - \bar{I}) \} \Delta(N \cos \bar{I}) \end{aligned}$$

To derive  $\sin V'$  from  $\sin V$  we use the invariance of the quantity

$$N \sin I \cos(\alpha - \bar{\alpha}) =$$

$$= N \sin I (\cos V \cos(I - \bar{I}) - \sin V \sin(I - \bar{I})).$$

To this we add and subtract

$N \cos I \sin V \cos(I - \bar{I})$  and have the invariance of

$$N \cos(I - \bar{I}) \sin(I + V) - N \cos \bar{I} \sin V.$$

Substituting from (iv)

$$(vi) \quad \begin{aligned} &\sin V' = \\ &\frac{N \cos \bar{I} \sin V - EK + E'K'}{N' \cos \bar{I}'}, \quad K \equiv N \cos(I - \bar{I})c \end{aligned}$$

(i) to (iv) with addition of the obvious results (1 and 2, denote consecutive surfaces)

$$(vii) \quad \bar{D} - D = \varepsilon'_1 - \varepsilon_2$$

$$(viii) \quad h_{r_2} = h'_{r_1} - \bar{D} \sin V$$

completes the equations needed to trace a T-ray.

In the degenerate case of an axial ray they reduce to

$$(ia) \quad h = Y + \varepsilon \sin U$$

$$(iia) \quad h = r (\sin I + \sin U)$$

$$(iiia) \quad \varepsilon = \frac{2 h \sin I}{G}, \quad G \equiv (1 + \cos U)(1 + \cos I) + \sin U \sin I$$

$$(viiia) \quad \sin U' = \frac{N \sin U - YK + YK'}{N'}, \quad K \equiv N \cos I c$$

$$(viia) \quad d - D = \varepsilon'_1 - \varepsilon_2$$

$$(viiiia) \quad h_2 = h'_1 - d \sin U$$

## REFERENCES

- CRUIKSHANK & MCAULEY, *Proc. Phys. Soc.*, **57**, 1945 Pts. 4 & 5, p. 350, 362, 419, 426, 430. Ibid. **58**, 1946, Pt. 3, p. 296.  
F. FREITAG, *Optik*, **9**, 1952, p. 529.  
ROSEN & ELDERT, *J. O. S. A.*, **44**, 1954, p. 250.  
F. A. LUCY, *J. O. S. A.*, **45**, 1955, p. 320, 323.  
N. J. RUMSEY, *J. O. S. A.*, **41**, 1951, p. 229.  
H. H. HOPKINS, *The Wave Theory of Aberrations*, 1950, Oxford University Press; *Proc. Phys. Soc.*, B, **65**, 1952, p. 934.  
W. S. S. BLASCHKE, *Proc. Phys. Soc.*, B, **67**, 1954, p. 801.

Manuscript reçu le 7 décembre 1955.

## The intensity at the surface of a semi-infinite diffuser with an infinite line source parallel to the surface

R. G. GIOVANELLI

Division of Physics, National Standards Laboratory,  
Commonwealth Scientific and Industrial Research  
Organization, Sydney.

**SUMMARY.** — Equations are developed for the intensity distribution at the surface of a semi-infinite isotropically scattering diffusing medium whose matrix may be of refractive index unity or greater, the source being a line lying on the surface; numerical solutions are given for several values of the albedo for single scattering and hence of the reflectance.

An extension of the argument leads to formal solutions for the intensity distribution for line sources below the surface.

**SOMMAIRE.** — On établit des relations qui donnent la distribution des intensités à la surface d'un milieu semi-infini doué d'une diffusion isotrope représentable par un modèle où l'indice est supérieur ou égal à l'unité; la source est une droite de la surface. Des solutions numériques sont fournies pour plusieurs valeurs de l'albedo (correspondant à une diffusion unique) et par conséquent de la réflectance.

Une généralisation du raisonnement conduit à des expressions de la distribution des intensités dans le cas où la source linéaire est au-dessous de la surface.

**ZUSAMMENFASSUNG.** — Es werden die Gleichungen für die Intensitätsverteilung an der Oberfläche eines isotrop streuenden Mittels abgeleitet, das den unendlichen Halbraum erfüllt und dessen Matrix den Brechwert 1 oder  $n$  hat. Die gradlinige Lichtquelle liegt auf der Oberfläche des Mediums. Die Zahlenergebnisse gelten für verschiedene Werte der Albedo des einzelnen Streuvorganges und damit auch der Rückstrahlung. Eine Ausdehnung der Untersuchung führt zu den Formelwerten für die Intensitätsverteilung bei einer linienförmigen Lichtquelle unter der Oberfläche.

**I. Introduction.** — While many problems involving radiative transfer have been solved for plane parallel media and for media possessing spherical or cylindrical symmetry, such discussions have been restricted to cases involving intensity variations in only one dimension. No solutions, however, have been given of problems such as arise when a finite source is located near the surface of a diffuser, or when a diffuser is illuminated non-uniformly.

In this paper we develop an approximate integral equation for the intensity distribution at the surface of a semi-infinite isotropically scattering diffuser for a line source on the surface; and from the solution of this equation give formal solutions of the surface intensity for line sources parallel to the surface but within the medium. The method depends on a knowledge of the intensity distribution for a line source in an infinite diffuser, recently derived by Jefferies [1] (1955).

The present paper contains the relevant analysis and numerical solutions for semi-infinite diffusers having various values of the albedo for single scattering.

**II. The intensity in an infinite isotropically scattering medium containing an infinite line source.** — The model assumed for the diffusing medium is a matrix (whose refractive index  $N$  may be greater than unity) containing randomly spaced centres which scatter radiation isotropically; radiation may also be absorbed by the medium. The coefficients of scattering and "true" absorption are denoted by  $\sigma$  and  $\alpha$  respectively, the extinction or attenuation coefficient

$\sigma + \alpha$  by  $\kappa$ ;  $\tilde{\omega}_0 = 1 - \lambda = \frac{\sigma}{\kappa}$  is the albedo for single scattering.

Let an infinite line source, emitting flux  $4\pi Q$  per unit length, be situated in an infinite diffusing medium. The direct flux from the source is converted into diffused flux by scattering, thus providing a source of diffused flux distributed through the medium. The source function  $B$ , which is the ratio of the diffused flux produced in unit volume and emitted into unit solid angle to the extinction coefficient, is given by

$$(2.1) \quad B = \frac{Q \tilde{\omega}_0}{2\pi} \int_0^\infty \frac{e^{-\kappa x}}{x \sqrt{x^2 - \rho^2}} dx.$$

JEFFERIES (1955) has shown that with an approximation analogous to that of EDDINGTON for a plane parallel medium, the equation of radiative transfer may be written in cylindrical co-ordinates

$$(2.2) \quad \frac{d^2 J}{d\rho^2} + \frac{1}{\rho} \frac{dJ}{d\rho} = 3\kappa^2(\lambda J - 4\pi B)$$

where  $\rho$  is the perpendicular distance from the source and  $J$  the total intensity of diffused radiation, defined as  $J = \int I d\Omega$  where the integration is carried out over all solid angles and  $I d\Omega$  is the radiant power crossing unit projected area in an element of solid angle. JEFFERIES has given the solution of (2.2) in the form of tables of

$$\frac{(J\rho)}{6\tilde{\omega}_0\kappa Q}$$

for selected values of  $\tilde{\omega}_0$ .



### III. The equation for the intensity at the surface of a semi-infinite diffuser having a matrix of refractive index unity and an infinite line source on the surface.

Consider an infinite line source on the surface of a semi-infinite isotropically scattering diffuser emitting in all directions flux  $4\pi Q$  per unit length, and let the total intensity of diffused light at the surface be  $Q\varphi(r)$ , where  $r$  is the distance from the source; at any point on the surface the diffused radiation is, of course, confined to outward directions. Then if two such semi-infinite media be placed in contact enclosing the line source, the intensity distribution will become that in an infinite medium (§ II), denoted by  $QF(r)$ . However, the latter may now be regarded as composed of (i) two components each independent of the presence of the other semi-infinite medium, totalling  $2Q\varphi(r)$ , and (ii) components due to the distributed or surface sources represented by the adjacent surface of the other medium.

Since the intensity of diffused radiation  $I$  is almost isotropic in an infinite medium, the flux passing from one of the above semi-infinite media to the other is closely  $\frac{J}{4}$  per unit area, so that there is effectively a surface source over the interface emitting a total flux of  $\frac{J}{2}$  per unit area. The surface source gives rise to a different polar distribution of "direct" light from that of the line source but the effect of this difference is disregarded for the present. The "direct" light itself contributes nothing to the total intensity along the plane separating the two media, for there the surface source subtends zero angle. Thus at distance  $R$  from the source we find the approximate equation

$$QF(R) = 2Q\varphi(R) + \frac{Q}{4\pi} \int_{-\infty}^{\infty} F(r) \varphi(R-r) dr$$

or

$$(3.1) \quad \varphi(R) = \frac{1}{2} F(R) - \frac{1}{8\pi} \int_{-\infty}^{\infty} F(r) \varphi(R-r) dr.$$

Since it is clear that  $\varphi(R) = \varphi(-R)$ , the above equation may be rewritten, with a slight change of variable,

$$(3.2) \quad \varphi(R) = \frac{1}{2} F(R) - \frac{1}{8\pi} \int_{-\infty}^{\infty} F(R+r) \varphi(r) dr.$$

**IV. The intensity in an infinite medium with an infinite opaque strip source of narrow width.** — The equations in § III have been derived neglecting differences in the intensity distribution of diffused radiation originating from line and from opaque strip sources. For a source of non-uniform polar distribution in an infinite medium the total intensity of diffused radiation may be found by replacing the source with an equivalent source distributed throughout the volume and representing once-scattered radiation. The total intensity of diffused radiation at a point is thus given by the sum of the total intensities of the

diffused and of the "direct" radiation from the distributed source integrated throughout space.

For an opaque strip source of narrow width emitting flux  $Q$  per unit length into an infinite medium it is not difficult to show in this way that at distance  $r$  from the source

$$(4.1) \quad J(r) = \frac{2Q\tilde{\omega}_0\kappa}{\pi} \int_0^{\infty} \int_{-1}^1 Ki_2(\kappa\rho) \left[ F(t) + \frac{2Ki_1(\kappa t)}{t} \right] d\mu d\rho$$

where  $t^2 = r^2 + \rho^2 - 2r\rho\mu$  for a point in a plane containing the strip, or  $t^2 = r^2 + \rho^2 - 2r\rho\sqrt{1-\mu^2}$  for a point in a plane at right angles to and containing the axis of the strip, and

$$Ki_n(x) = \int_0^{\infty} \frac{e^{-x} \cosh u}{\cosh^n u} du.$$

The integrand in (4.1) is very small when  $\kappa\rho \gg 1$ ; thus the integral over  $\rho$  need not be extended beyond some value  $\rho_1 \gg \kappa^{-1}$ . If  $r \gg \rho_1$ , the values of  $t$  and hence of  $J(r)$  are the same in both planes. It follows that at sufficient distances the intensity is independent of the polar distribution of direct radiation from the source, and this will probably be so at distances for which JEFFERIES' solution of

$$\frac{J(\rho)}{6\pi_0\kappa Q}$$

approaches its asymptotic value

$$\frac{\beta}{2} \log \left( \frac{\beta+1}{\beta-1} \right) K_0 \left( \frac{\kappa\rho}{\beta} \right)$$

where  $\beta = (3\lambda)^{-1/2}$ .

At distances from the source small compared with  $\kappa^{-1}$ , the total intensity of diffused radiation is also clearly independent of the polar distribution of direct radiation from the source. By inference, the variations at intermediate distances must also be small, particularly for the larger values of  $\beta$ . Reference to JEFFERIES' solution suggests that for  $\beta = 5$  i.e.,  $\tilde{\omega}_0 = \frac{74}{75}$ , the differences for different types of source are probably not more than about 1 per cent, and are even smaller for  $\beta = 10$ ; for  $\beta = 2$  they may be appreciably greater.

In view of the labour involved in evaluating (4.1), and the small differences expected, the results are assumed to be the same as for a line source for the purposes of the present discussion.

**V. The equation for the intensity at the surface of a semi-infinite diffuser having a matrix of refractive index  $N$  and a line source on the surface.** — Consider a semi-infinite diffuser with a line source lying in optical contact with the surface, and emitting into the medium flux  $2\pi Q$  per unit length. At distance  $r$  from the

source, and just within the surface, let the contribution to the total intensity from radiation directed towards the surface be  $Q \varphi(r)$ ; thus the flux incident on the surface from below is very closely  $\frac{1}{2} Q \varphi(r)$  per unit area. Since the transmission of the surface to uniformly diffused radiation incident from within is  $\mathfrak{J}/N^2$ , where  $\mathfrak{J}$  is the transmission of the surface to uniformly diffused radiation incident from air, the escaping flux is very closely  $\frac{1}{2} Q \varphi(r) (\mathfrak{J}/N^2)$  per unit area.

Now let two such semi-infinite media be placed almost in contact, the source emitting flux  $2 \pi Q$  per unit length into each medium. The distribution  $QF(r)$  is thus restored. As before, this may be regarded as composed of (i) components independent of the presence of the other semi-infinite medium, and (ii) components due to the distributed or surface source represented by the adjacent surface of the other medium. In the case of (i), radiation reflected at the surface and transmitted through it are complementary, so that the total intensity due to these components is  $2 Q \varphi(r)$ . The surface source due to radiation transmitted from one medium to the other effectively emits a total flux of  $\frac{1}{2} Q (\mathfrak{J}/N^2) F(r)$  per unit area, this including the effects of multiple reflections between the surfaces. Any element of this effective plane source has a polar distribution of radiation which depends on the glossiness or optical state of the surface but, as before, it is sufficient to assume that the total intensity of diffused radiation due to a source of this type is the same as for a distribution of line sources. Thus at distance  $R$ ,

$$Q F(R) = 2 Q \varphi(R) + \frac{Q \mathfrak{J}}{4 \pi N^2} \int_{-\infty}^{\infty} F(r) \varphi(R-r) dr$$

or

$$(5.1) \quad \varphi(R) = \frac{1}{2} F(R) - \frac{\mathfrak{J}}{4 \pi N^2} \int_{-\infty}^{\infty} F(R+r) \varphi(r) dr.$$

This equation reduces to (3.2) when  $N = 1$ .

**VI. Integral properties of  $\varphi(r)$  and  $F(r)$ .** — Since when  $N = 1$  the escaping flux is very closely  $\frac{1}{2} \pi$  per unit area, it follows that for diffused flux  $2 \pi Q$  incident on the medium the total reflected flux is

$$\frac{Q}{2} \int_{-\infty}^{\infty} \varphi(r) dr.$$

The ratio of the latter to the former is the reflectance for diffused incident radiation,  $\mathfrak{R}$  so that

$$(6.1) \quad \int_{-\infty}^{\infty} \varphi(r) dr = 4 \pi \mathfrak{R}.$$

Integrating (3.1) with respect to  $R$  over the range  $-\infty$  to  $\infty$

$$\int_{-\infty}^{\infty} \varphi(R) dR = \frac{1}{2} \int_{-\infty}^{\infty} F(R) dR - \frac{1}{8 \pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(r) \varphi(R-r) dr dR.$$

Carrying out the double integration over  $R$  first, it follows from (6.1) that

$$(6.2) \quad \int_{-\infty}^{\infty} F(R) dR = \frac{8 \pi \mathfrak{R}}{1 - \mathfrak{R}}.$$

When  $N > 1$ , it may be shown that

$$(6.3) \quad \int_{-\infty}^{\infty} \varphi(r) dr = \frac{4 \pi N^2}{\mathfrak{J}^2} \mathfrak{R},$$

where  $\mathfrak{R}$  is now the reflectance for diffused incident radiation omitting FRESNEL reflectance, or

$$\mathfrak{R} = \mathfrak{R}_D - (1 - \mathfrak{J}),$$

where  $\mathfrak{R}_D$  is the total reflectance for diffused incident radiation.

**VII. Solution of the intensity equation.** — Equation (3.2) may be recast slightly by noting that  $\varphi(r)$  may be expected to decrease with increasing  $r$  more rapidly than does  $F(r)$ . Thus

$$\chi(r) = \frac{\varphi(r)}{F(r)}$$

is expected to vary less rapidly than  $\varphi(r)$ . Equation (3.2) may then be rewritten as an equation in  $\chi(r)$ :

$$(7.1) \quad \chi(R) = \frac{1}{2} - \frac{1}{8 \pi F(R)} \int_{-\infty}^{\infty} F(R+r) F(r) \chi(r) dr.$$

The function  $\chi(r)$  is now replaced by a polynomial in  $r$ ; or alternatively  $\chi(r)$  is allocated unknown values  $\alpha, \beta, \gamma, \delta, \dots$  at equally spaced intervals of  $r$ , intermediate values being obtained in terms of these by Lagrangian interpolation. For a given value of  $\varpi_0$ ,  $F(R)$  is obtained from JEFFERIES' tables and by noting that for large values of  $R$ ,

$$\frac{F(R)}{6 \varpi_0 \kappa} = \frac{\beta}{2} \log \left( \frac{\beta + 1}{\beta - 1} \right) K_0 (\sqrt{3 \lambda} \kappa R),$$

where  $\beta = (3 \lambda)^{-1/2}$ ; graphical interpolation for small values of  $R$  enables values of  $F(R)$  to be obtained with sufficient accuracy at close intervals. The integral in (7.1) can thus be evaluated numerically as a linear function of  $\alpha, \beta, \gamma, \delta, \dots$ , for given values of  $R$ ; for the latter, a choice of the same values as for  $\alpha, \beta$ , etc., enables (7.1) to be replaced by a set of linear simultaneous equations for  $\alpha, \beta, \gamma, \delta, \dots$ , which can be solved in the usual way.



In this method of solution, two points need special mention ; (i) it is desirable to replace the variable  $r$  in the integral of (7.1) by  $y = \sqrt{3 \lambda \kappa} r$ , in terms of which  $F(R)$  is more readily expressed ; (ii) rather than evaluate the integral for values of  $r$  for which  $\chi(r)$  would need to be extrapolated, the contribution for these larger values of  $r$  (which will be small in any case) is found by assuming an exponential decrease in the integrand, whose rate is estimated from the ratio  $\delta/\gamma$  obtained by trial solutions.

The solutions given in Tables 1-3 have been obtained with four values of  $\chi(r)$ , the corresponding values of  $y$  and  $\frac{\varphi(r)}{\kappa}$  being indicated in the appropriate table, while graphs of  $\frac{\varphi(r)}{\kappa}$  in figure 1 represent the functions derived using four-point Lagrangian interpolation between the  $\chi(r)$ .

At no stage can the accuracy of these results be expected to be better than one per cent, and it may be several times poorer. The EDDINGTON approximation

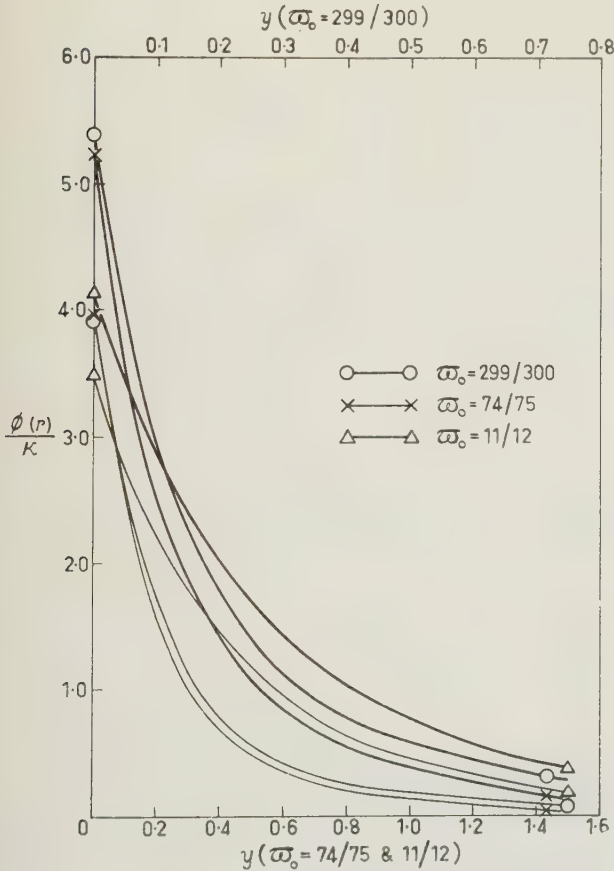


FIG. 1. — The function  $\frac{\varphi(r)}{\kappa}$ . The values of the abscissae,  $y = \sqrt{3 \lambda \kappa} r$ , depend on  $\bar{\omega}_0$ , those for  $\bar{\omega}_0 = \frac{299}{300}$  being indicated at the top, and for  $\bar{\omega}_0 = \frac{74}{75}$  and  $\frac{11}{12}$  at the bottom, of the graph. Light and heavy curves are for  $N = 1$  and  $N = 1.5$  respectively.

TABLE 1

Intensity distribution at the surface of a semi-infinite isotropic diffuser with a line source on the surface

$$\bar{\omega}_0 = \frac{299}{300} = 0.996667$$

		$y$			
		0	0.25	0.5	0.75
$N = 1$	$\frac{F(r)}{6 \bar{\omega}_0 \kappa}$	3.148	1.541	0.927	0.613
	$\chi(r)$	0.2074	0.0573	0.0323	0.0176
	$\frac{1}{\kappa} \varphi(r)$	3.904	0.528	0.179	0.064
$N = 1.5$	$\chi(r)$	0.2865	0.1493	0.1012	0.0728
	$\frac{1}{\kappa} \varphi(r)$	5.393	1.376	0.561	0.267

TABLE 2

Intensity distribution at the surface of a semi-infinite isotropic diffuser with a line source on the surface

$$\bar{\omega}_0 = \frac{74}{75} = 0.986667$$

		$y$			
		0	0.5	1.0	1.5
$N = 1$	$\frac{F(r)}{6 \bar{\omega}_0 \kappa}$	2.554	0.885	0.427	0.217
	$\chi(r)$	0.2623	0.0879	0.0510	0.0266
	$\frac{1}{\kappa} \varphi(r)$	3.966	0.460	0.129	0.034
$N = 1.5$	$\chi(r)$	0.3462	0.2029	0.1485	0.1067
	$\frac{1}{\kappa} \varphi(r)$	5.234	1.063	0.375	0.137

TABLE 3

Intensity distribution at the surface of a semi-infinite isotropic diffuser with a line source on the surface

$$\bar{\omega}_0 = \frac{11}{12} = 0.916667$$

		$y$			
		0	0.5	1.0	1.5
$N = 1$	$\frac{F(r)}{6 \bar{\omega}_0 \kappa}$	1.792	0.843	0.430	0.234
	$\chi(r)$	0.3533	0.2494	0.1863	0.1462
	$\frac{1}{\kappa} \varphi(r)$	3.482	1.156	0.441	0.188
$N = 1.5$	$\chi(r)$	0.4207	0.3598	0.3160	0.2816
	$\frac{1}{\kappa} \varphi(r)$	4.146	1.669	0.747	0.362

is known to introduce errors of this magnitude, and JEFFERIES' equation of transfer probably introduces comparable errors. Further a small error has been introduced in deriving (3.1) by assuming that the line source and the surface source produce intensity

distributions along the plane of separation of the two media which are of identical form  $\varphi(r)$ . Finally, the solution depends on the replacement of  $\chi(r)$  by a polynomial, and this necessarily introduces further errors, the effect of which may be noted in the computed value of  $\frac{\varphi(r)}{x}$  at the source, which is nearly 2 per cent larger for  $\bar{\omega}_0 = \frac{74}{75}$  than for  $\bar{\omega}_0 = \frac{299}{300}$ , whereas it should be smaller. However, extremely useful checks can be made via the integral relations (6.1) and (6.2). The numerical value of  $\int_{-\infty}^{\infty} F(r) dr$  obtained from JEFFERIES' solution agrees to within one per cent with the value  $\frac{4\pi R}{1-R}$  (derived neglecting the differences in polar distribution of the sources) for the two higher values of  $\bar{\omega}_0$ , but is rather higher for the lowest value of  $\bar{\omega}_0$ . Further the integral  $\int_{-\infty}^{\infty} \varphi(r) dr$  evaluated numerically from our solutions agrees satisfactorily with the value  $4\pi R$ , as may be seen from Table 4.

TABLE 4

The Integrals  $\int_0^{\infty} F(r) dr$  and  $\int_{-\infty}^{\infty} \varphi(r) dr$ ;  $N = 1$

$\bar{\omega}_0$	$\frac{11}{12}$	$\frac{74}{75}$	$\frac{299}{300}$
$R \left\{ \begin{array}{l} \text{Exact} \dots\dots\dots \\ \text{On EDDINGTON approx.} \end{array} \right.$	$\begin{array}{l} 0.5111 \\ 0.5199 \end{array}$	$\begin{array}{l} 0.7668 \\ 0.7695 \end{array}$	$\begin{array}{l} 0.8756 \\ 0.8764 \end{array}$
$\int_0^{\infty} F(r) dr \dots\dots\dots$	14.11	41.7	89.6
$\frac{4\pi R}{1-R} \left\{ \begin{array}{l} \text{Exact} \dots\dots\dots \\ \text{On EDDINGTON approx.} \end{array} \right.$	$\begin{array}{l} 13.14 \\ 13.61 \end{array}$	$\begin{array}{l} 41.3 \\ 42.0 \end{array}$	$\begin{array}{l} 88.4 \\ 89.1 \end{array}$
$\int_{-\infty}^{\infty} \varphi(r) dr \dots\dots\dots$	6.65	9.60	10.77
$4\pi R \left\{ \begin{array}{l} \text{Exact} \dots\dots\dots \\ \text{On EDDINGTON approx.} \end{array} \right.$	$\begin{array}{l} 6.42 \\ 6.53 \end{array}$	$\begin{array}{l} 9.64 \\ 9.67 \end{array}$	$\begin{array}{l} 11.00 \\ 11.01 \end{array}$

VIII. The intensity distribution at the surface of a semi-infinite diffuser with an infinite line source beneath the surface. — For an infinite line source at distance  $Z$  beneath the surface of a semi-infinite diffuser, and emitting flux  $4\pi Q$  per unit length, let the total intensity of diffused radiation at the surface be  $Q$ .  $\theta(r)$  at distance  $r$  from the projection of the source on the surface.

For a matrix of refractive index unity, suppose as before that a similar semi-infinite diffuser is now placed in contact with the surface, the intensity distribution of the diffused radiation being thus increased to  $QF(\rho)$ , where  $\rho$  is the distance from the line source. At the interface this total intensity is composed of

(i) the independent component due to the first medium (ii) the component in each medium due to the distributed or surface source represented by the adjacent surface of the other medium, and (iii) the component in the second medium due to irradiation by residual direct light from the source, which would normally have escaped from the first medium. The last component is somewhat different in character from those discussed previously, and will now be evaluated.

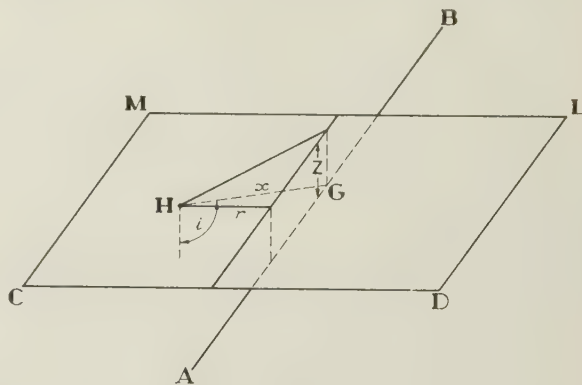


FIG. 2

FIG. 2. — Line source AB beneath the surface of a semi-infinite diffuser C D L M.

Let AB (fig. 2) be the line source at distance  $Z$  below the surface CDLM of the diffuser. Then the surface source due to direct radiation on the upper medium can be evaluated by finding the direct illumination on the surface CDLM. From an element of source of length  $dl$  at G, the illumination  $dE$  at H, distant  $r$  from the projection of the source on the surface, is

$$dE = \frac{Q dl \cdot e^{-x} \cos i}{x^2}$$

where  $x = GH$  is the distance of the element of source from the point under consideration, and  $i$  is the angle GH makes with the normal to the surface. Thus

$$E = \int_{-\infty}^{\infty} \frac{Q e^{-x} \cos i}{x^2} dl.$$

But  $\cos i = \frac{Z}{x}$  and  $ldl = x dx$ , so

$$(8.1) \quad E(r) = 2 Q Z \int_{\rho}^{\infty} \frac{e^{-x}}{x^2 \sqrt{x^2 - Z^2 - r^2}} dx.$$

This is the flux entering unit area of the upper medium: assuming here too that the surface source is associated with a  $\varphi(r)$  of the same form as the line source, at distance  $R$  from the projection of the source on the surface it gives rise in the upper medium to a component of the total intensity of diffused radiation

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} E(r) \varphi(R - r) dr.$$

We may thus proceed to write down the approximate equation for  $\theta(R)$ :



$$(8.2) \quad Q F(\sqrt{R^2 + Z^2}) = \\ = Q\theta(R) + \frac{Q}{4\pi} \int_{-\infty}^{\infty} F(\sqrt{r^2 + Z^2}) \varphi(R-r) dr + \\ + \frac{1}{2\pi} \int_{-\infty}^{\infty} E(r) \varphi(R-r) dr.$$

The formal solution for  $\theta(R)$  follows immediately :

$$(8.3) \quad \theta(R) = F(\sqrt{R^2 + Z^2}) - \\ - \frac{1}{4\pi} \int_{-\infty}^{\infty} F(\sqrt{r^2 + Z^2}) \varphi(R-r) dr - \\ - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{E(r)}{Q} \varphi(R-r) dr.$$

For sources well below the surface, so that  $e^{-\kappa Z}$  is negligible,  $E(r)$  is negligible and (8.3) simplifies to

$$(8.4) \quad \theta(R) = \\ F(\sqrt{R^2 + Z^2}) - \frac{1}{4\pi} \int_{-\infty}^{\infty} F(\sqrt{r^2 + Z^2}) \varphi(R-r) dr.$$

It must be remembered that (8.3) represents only the diffused radiation. The total radiation field is found by adding the direct component.

For a matrix of refractive index  $N$ , suppose that an identical semi-infinite medium containing a similar line source is placed almost in contact with the first medium, the line joining the sources being perpendicular to the surface. Then the total diffused intensity at any point on the surface distant  $R$  from the line joining the sources is  $2 Q F(\sqrt{R^2 + Z^2})$ . This is composed of

(i) two independent components totalling  $2 Q\theta(R)$ , where  $Q\theta(R)$  is the total outward intensity in a hemisphere just below the surface when only one semi-infinite medium is present ;

(ii) two additional components totalling

$$\frac{2}{2\pi} \int_{-\infty}^{\infty} E(r) t(r) \varphi(R-r) dr,$$

arising from direct radiation from the one source incident on the other medium,  $t(r)$  being the transmittance of a single surface of refractive index  $N$  for " direct " light from this source — in the case of a polished surface,  $t(r)$  being given by FRESNEL's well-known rules ; and

(iii) components due to the distributed or surface source represented by the adjacent surface of the other medium, and totalling

$$\frac{2 Q \mathfrak{J}}{4\pi N^2} \int_{-\infty}^{\infty} F(\sqrt{r^2 + Z^2}) \varphi(R-r) dr$$

(see § IV).

Thus

$$2 Q F(\sqrt{R^2 + Z^2}) = \\ = 2 Q \theta(R) + \frac{Q \mathfrak{J}}{2\pi N^2} \int_{-\infty}^{\infty} F(\sqrt{r^2 + Z^2}) \varphi(R-r) dr + \\ + \frac{1}{\pi} \int_{-\infty}^{\infty} E(r) t(r) \theta(R-r) dr.$$

The formal solution for  $\theta(R)$  follows :

$$(8.5) \quad \theta(R) = F(\sqrt{R^2 + Z^2}) - \\ - \frac{\mathfrak{J}}{4\pi N^2} \int_{-\infty}^{\infty} F(\sqrt{r^2 + Z^2}) \varphi(R-r) dr - \\ - \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{E(r)}{Q} t(r) \varphi(R-r) dr.$$

For sources well below the surface, so that  $e^{-\kappa Z}$  is negligible,  $E(r)$  is negligible and (8.5) simplifies to

$$(8.6) \quad \theta(R) = F(\sqrt{R^2 + Z^2}) - \\ - \frac{\mathfrak{J}}{4\pi N^2} \int_{-\infty}^{\infty} F(\sqrt{r^2 + Z^2}) \varphi(R-r) dr.$$

Again,  $\theta(R)$  is only the diffused component of the outward radiation field immediately below the surface ; the total diffused intensity is  $\theta(R) \left(2 - \frac{\mathfrak{J}}{N^2}\right)$  ; the total radiation field is found by adding the direct component.

**IX. Discussion.** — It has been pointed out that the function  $\varphi(r)$  depends on the polar distribution of direct radiation from the source. We have shown that as  $\tilde{\omega}_0$  approaches unity the solution becomes almost independent of the polar distribution, so that the results given here are good approximations for problems such as the distribution of intensity along the surface of a semi-infinite diffuser for light incident from any cone of illumination. The resolving power of the surface used as a projection screen may be assessed from the range  $\pm \bar{Y}$  in which half the light is reflected. In Table 5 this is given in terms of  $\bar{Y} = \sqrt{3 \lambda \kappa R}$ .

TABLE 5  
Range  $\pm \bar{Y}$  in which half the light is reflected

$\tilde{\omega}_0$	$\bar{Y}$	
	$N = 1$	$N = 1.5$
299	0.10	0.17
300		
74	0.18	0.26
75		
11	0.35	0.40
12		

The above results also give, with appropriate changes of terminology, the distribution of thermal neutrons along the surface of a semi-infinite medium with a line source of thermal neutrons parallel to the surface ; for the diffusion equations governing radiation and thermal neutrons are formally identical.

**X. Acknowledgment.** — The author wishes to thank Mr J. T. JEFFERIES for valuable discussion on the above subject, and for making available his solutions for an infinite diffuser prior to publication.

#### REFERENCE

- [1] J. T. JEFFERIES, *Opt. Acta* 2, 1955, p. 109.

*Manuscript reçu le 20 juillet 1955.*

## Gekoppelte Interferenzfilter

ULRICH GRADMANN

Institut für angewandte Physik, Marburg.

**ZUSAMMENFASSUNG.** — Einfache Interferenzfilter bestehen aus zwei parallelen, in Dielektrika eingebetteten Metallschichten. Geht man zu Systemen mit mehr als zwei parallelen Metallschichten über, setzt also mehrere solche Einfachfilter hintereinander, so multiplizieren sich deren Durchlässigkeiten nicht einfach, es treten vielmehr Kopplungseffekte auf. Für Filter mit 2, 3 und 4 Metallschichten werden Formeln für den spektralen Verlauf der Durchlässigkeiten angegeben und diskutiert. Numerisch gerechnete Beispiele werden in Diagrammen dargestellt.

**SOMMAIRE.** — Les filtres interférentiels simples se composent de deux couches métalliques parallèles entourées par des diélectriques. Si l'on passe à des systèmes comprenant plus de deux couches métalliques parallèles, c'est-à-dire lorsqu'on met en série plusieurs filtres simples leurs intensités ne se multiplient pas simplement; de plus, des effets de couplage apparaissent. On donne et discute les formules pour la variation spectrale des facteurs de transmission pour des filtres à 2, 3 et 4 couches métalliques. Des exemples calculés numériquement sont représentés sur des diagrammes.

**SUMMARY.** — Simple interference filters consist of two parallel metallic layers alternating with dielectrics. For systems of more than two parallel metallic layers, i. e. when several simple filters are placed in series, their transmissions are not simply multiplied, but coupling effects appear. Formulae are given and discussed for the spectral variation of the transmission factors for filters of 2, 3 and 4 metallic layers. Diagrams are given of calculated numerical examples.

1. **Einleitung.** — Reflexion und Durchlässigkeit eines einfachen Interferenzfilters aus zwei dünnen Metallschichten und einem dazwischen liegenden Dielektrikum sind theoretisch seit langem geklärt. In seiner Patentschrift zu diesem Thema zeigt GEFFCKEN [1], wie man durch Hintereinanderschalten entkoppelter, d. h. unter einem spitzen Winkel zueinander stehender Elementarfilter zu Filtersystemen mit sehr scharfen Linien kommen kann. Werden die Filter jedoch nicht entkoppelt, stehen sie also parallel, so treten kompliziertere Verhältnisse auf, die hier durch Zurückgehen auf die MAXWELL'sche Theorie diskutiert werden sollen. Das Problem steht in Analogie zur Vierpoltheorie der Hochfrequenzschaltungen, worauf SCHUSTER [2] hingewiesen hat, und man erhält dementsprechend auch hier Bandfilter, d. h. Filter mit einem verbreiterten Durchlassbereich hoher Flankensteilheit <sup>(1)</sup> [3].

Unter einigen vereinfachenden Voraussetzungen, die das Wesentliche klar herausstellen, sollen nun solche gekoppelten Filter besprochen und Zahlenbeispiele angegeben werden. Wir beschränken uns auf senkrechten Einfall, nehmen alle Metallschichten als gleich an, betten sie in ein einheitliches Dielektrikum ein und vernachlässigen die Dispersion der optischen Konstanten.

2. **Grundlagen.** — Ein System paralleler Schichten sei gegeben, bestehend aus den Teilsystemen I und II, als ganzes eingebettet in ein Dielektrikum vom Brechungsindex  $n_0$ . Wir betrachten nun eine senkrecht einfallende, der  $z$ -Richtung entgegenlaufende, eben und linear polarisierte Welle (Abb. 1).

$$\mathcal{E} = \mathcal{E}_0 e^{i\omega\left(t + \frac{n_0 z}{c}\right)}$$

Die am Gesamtsystem reflektierte Welle sei

$$\mathcal{R} = \mathcal{R}_0 e^{i\omega\left(t - \frac{n_0 z}{c}\right)}$$

und die durchgehende

$$\mathcal{D} = \mathcal{D}_0 e^{i\omega\left(t + \frac{n_0 z}{c}\right)}$$

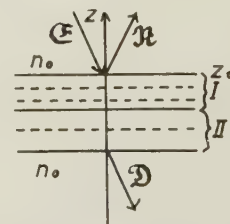


Abb. 1. — Ein System paralleler Schichten, bestehend aus den Teilsystemen I und II, ist in ein Dielektrikum vom Brechungsindex  $n_0$  eingebettet. Einfallende ( $\mathcal{E}$ ), reflektierte ( $\mathcal{R}$ ) und durchgehende Welle ( $\mathcal{D}$ ).

Wir setzen dann

$$(1a) \quad \frac{\mathcal{D}_0}{\mathcal{E}_0} = \delta^0 \quad \text{und} \quad (1b) \quad \frac{\mathcal{R}_0}{\mathcal{E}_0} = r^0$$

und analog für senkrechten Einfall von unten, also mit der  $z$ -Richtung,  $\frac{\mathcal{D}'_0}{\mathcal{E}'_0} = \delta^{0'}$  und  $\frac{\mathcal{R}'_0}{\mathcal{E}'_0} = r^{0'}$ ,

ferner für die Teilsysteme im gleichen Sinne  $\delta^0_I$ ,  $\delta^0_{II}$  usw. Dabei soll die Null als oberer Index bedeuten, dass alle Wellen auf den Nullpunkt desselben Koordinatensystems bezogen sind. Dann gilt <sup>(2)</sup>

$$(2) \quad \delta^0 = \frac{\delta^0_I \delta^0_{II}}{1 - r^{0'}_I r^0_{II}}$$

$$(3) \quad r^0 = \frac{r^0_I + r^0_{II} (\delta^0_I \delta^{0'}_I - r^0_I r^{0'}_I)}{1 - r^{0'}_I r^0_{II}}$$

3. **Das Einfachfilter.** — Das in diesem Abschnitt behandelte Einfachfilter besteht aus zwei Metallschichten der Dicke  $h$ , vom Brechungsindex  $n$ , die im

<sup>(1)</sup> Inzwischen wurden solche Filter auch experimentell verwirklicht [3].

<sup>(2)</sup> Siehe z. B. den Artikel « Optik dünner Schichten » in der Neuauflage des Handbuchs der Physik (im Druck).



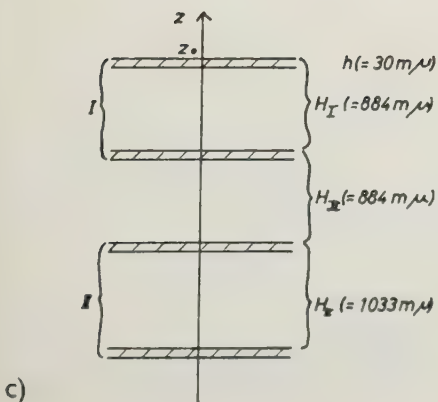
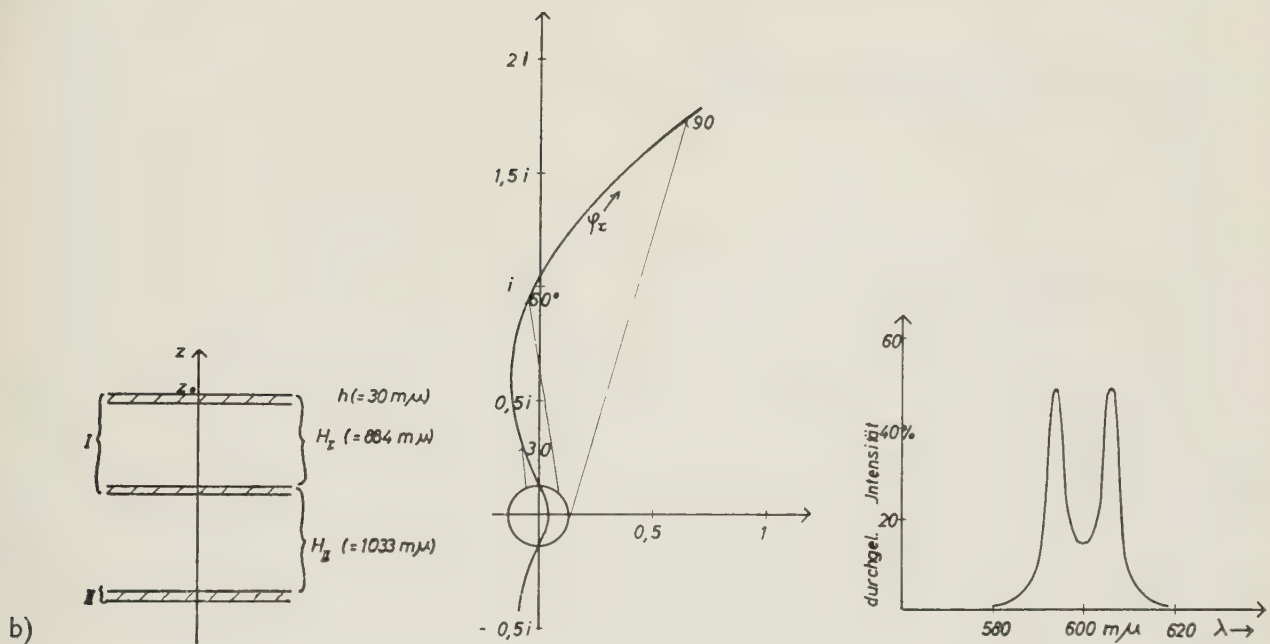
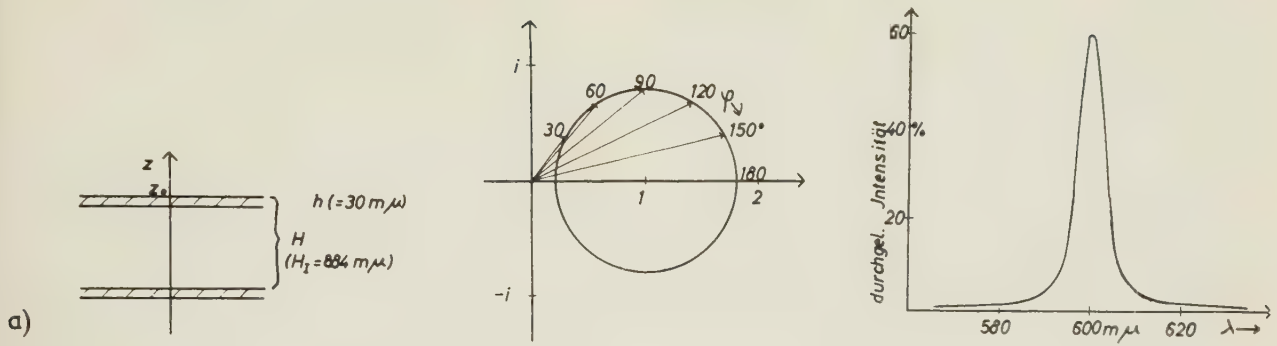
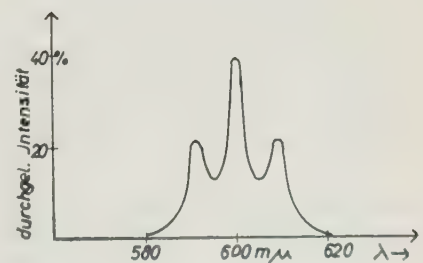


ABB. 2. — Verschiedene Filter.  
a) Einfachfilter, b) Doppelfilter, c) Dreifachfilter.

Erste Spalte: Aufbau der Filter.  
Zweite Spalte: Nenner der Durchlässigkeiten  $_{1D}$  (a) bzw.  $_{2D}$  (b) in der Zahlenebene. Die offene Kurve in (b) ist der erste Nennersummand von  $_{2D}$ , der Kreis das Kopplungsglied, die Pfeile sind der Nenner selbst. Parameter  $\varphi$  bzw.  $\varphi_1$  in Grad.

Dritte Spalte: Durchgelassene Intensitäten in % der eingefallenen.



Abstand  $H$  in ein Dielektrikum vom Brechungsindex  $n_0$  eingebettet sind. (Abb. 2 a).

Zunächst seien Reflexion  $_{or}$  (1) und Durchlässigkeit  $_{od}$  einer einzelnen Metallschicht berechnet, die im Sinne von Gl. (1) definiert und auf  $z_0 = 0$  bezogen seien. Man geht dazu mit den bekannten FRESNEL'schen Koeffizienten  $r$  und  $d$  in (2) und (3) ein (die "Systeme" I und II sind hier die beiden Grenzflächen), beachtet aber dabei, dass diese Koeffizienten auf die jeweiligen Grenzflächen selbst bezogen sind. Sie stehen zu unseren Grössen also in den folgenden Beziehungen (Abb. 3 b) :

$$d = \frac{2n_0}{n_0 + n_1} = \frac{\mathfrak{D}(z_1)}{\mathfrak{E}(z_1)} = \frac{\mathfrak{D}_0 e^{\rho_1 z_1}}{\mathfrak{E}_0 e^{\rho_0 z_1}} = d^0 e^{(\rho_1 - \rho_0)z_1}$$

$$r = \frac{n_0 - n_1}{n_0 + n_1} = \frac{\mathfrak{R}(z_1)}{\mathfrak{E}(z_1)} = \frac{\mathfrak{R}_0 e^{-\rho_0 z_1}}{\mathfrak{E}_0 e^{\rho_0 z_1}} = r^0 e^{-2\rho_0 z_1}$$

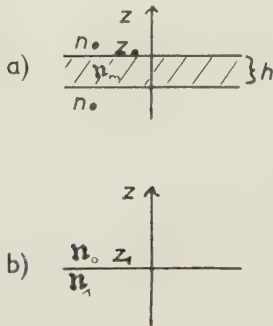


ABB. 3. — a) Eine Metallschicht der Dicke  $h$  vom Brechungsindex  $n_m$  in ein Dielektrikum vom Brechungsindex  $n_0$  eingebettet. b) Grenzfläche zweier Medien mit den Brechungsindizes  $n_0$  und  $n_1$ .

Damit erhält man

$$(4) \quad _{od} = \frac{\frac{4n_0 n_m}{(n_0 + n_m)^2} e^{-\rho_m h}}{1 - \left( \frac{n_0 - n_m}{n_0 + n_m} \right)^2 e^{-2\rho_m h}} e^{\rho_0 h} = _{od}^* e^{\rho_0 h}$$

$$(5) \quad _{or} = \frac{n_0 - n_m}{n_0 + n_m} \frac{1 - e^{-2\rho_m h}}{1 - \left( \frac{n_0 - n_m}{n_0 + n_m} \right)^2 e^{-2\rho_m h}}$$

wobei

$$(6) \quad \rho_i = \frac{2\pi i n_i}{\lambda} \text{ und } \lambda \text{ die Vakuumwellenlänge ist.}$$

Im Rechenbeispiel ist die Schicht als  $30 \text{ m}\mu$  dicke Silberschicht, ferner  $n_0 = 2$  angenommen. Der Zahlenwert für  $n_m$  nach KRAUTKRÄMER, MURMANN, WOLTER wurde dem Buche von H. MAYER [4] über dünne Schichten entnommen. Für  $h = 30 \text{ m}\mu$ ,  $\lambda = 600 \text{ m}\mu$  findet man dort

(1) Der Index wurde vorgestellt, um Konflikte mit den Bezeichnungen des oben zitierten Handbuchartikels zu vermeiden.

$$n_m = 0,12 - 4,15 i.$$

Setzt man diesen Wert in (4) und (5) ein so erhält man

$$(4a) \quad _{od}^* = 0,405 e^{0,526 i}$$

$$(5a) \quad _{or} = 0,888 e^{2,180 i}$$

Dass sich  $_{od}^*$  und  $_{or}$  in der Phase um nahezu  $90^\circ$  unterscheiden, folgt aus (4) und (5) unter Beachtung von  $\varphi_{n_m} \approx -\frac{\pi}{2}$  und wird weiter unten benutzt ( $\varphi_r$  sei die Phase einer komplexen Zahl  $r$ ).

$_{or}$  und  $_{od}$  zeigen an sich natürlich Dispersion, wurden den weiteren Ausführungen jedoch als Konstanten zugrunde gelegt, um das Grundsätzliche klarer herauszustellen.

Um nun Reflexion  $_{1r}$  und Durchlässigkeit  $_{1d}$  des Einfachfilters zu bekommen, die im Sinne von Gl. (1) auf  $z_0 = 0$  bezogen seien (Abb. 2 a), hat man nur mit (4) und (5) in (2) und (3) einzugehen, wobei freilich die Koordinatentransformation für die zweite Metallschicht zu beachten ist. Man erhält

$$_{1d} = \frac{(_{od})^2}{1 - _{or}' _{or} e^{2\rho_0 z_1}} = \frac{(_{od})^2}{1 - (_{or})^2 e^{-2\rho_0(H-h)}}$$

$$_{1r} = _{or} \frac{1 - e^{-2\rho_0 H} \{ (_{or})^2 e^{2\rho_0 h} - (_{od})^2 \}}{1 - _{or}' _{or} e^{2\rho_0 z_1}}$$

$$= _{or} \frac{1 - e^{-2\rho_0(H-h)} \{ (_{or})^2 - (_{od}^*)^2 \}}{1 - (_{or})^2 e^{-2\rho_0(H-h)}}$$

wenn man  $_{or}' = _{or} e^{2\rho_0 h}$  beachtet. Mit

$$(6) \quad \frac{2\rho_0}{i} (H - h) - 2\varphi_r = \varphi$$

und wenn man annimmt, dass sich  $_{od}^*$  und  $_{or}$  um  $90^\circ$  in der Phase unterscheiden, was, wie oben bemerkt, hier mit hinreichender Näherung zutrifft, wird aus den letzten Gleichungen

$$(7) \quad _{1d} = \frac{(_{od})^2}{1 - |_{or}|^2 e^{-i\varphi}}$$

$$(8) \quad _{1r} = _{or} \frac{1 - e^{-i\varphi} \{ |_{or}|^2 + |_{od}|^2 \}}{1 - |_{or}|^2 e^{-i\varphi}}$$

Da  $_{od}$  und  $_{or}$  Konstanten sein sollen, stellt der Nenner von (7) und (8) einen Kreis dar, wie in Abb. 2a zu sehen ist, aus der man auch deutlich sieht, wie die Maxima für ein Durchlässigkeitsfilter umso schärfer sind, je mehr sich  $_{or}$  dem Betrage nach 1 nähert. Daher wurde die Silberschicht  $30 \text{ m}\mu$  stark gewählt, was einigermassen scharfe Maxima ergab.

Als Spiegelung des Nenners am Kreise  $|r| = |_{od}|$  stellt auch  $_{1d}$  einen Kreis dar. Dasselbe gilt auch für  $_{1r}$ , wie man leicht nachrechnen kann. Abb. 4a und 4b zeigen  $_{1r}$  und  $_{1d}$  in der Zahlenebene.

$_{1r}$  und  $_{1d}$  sind in  $2\pi$  periodische Funktionen von



$$\varphi = \frac{2\rho_0}{1} (H-h) - 2\varphi_0 r = \frac{4\pi n_0}{\lambda} (H-h) - 2\varphi_0 r = \frac{4\pi n_0}{c} \nu (H-h) - 2\varphi_0 r$$

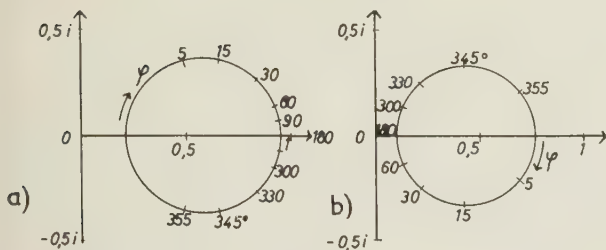


ABB. 4. — a) Reflexion  $1r$  und b) Durchlässigkeit  $1d$  des Einfachfilters in der Zahlenebene.

und damit auch periodische Funktionen der Frequenz  $\nu$ .

Maxima der Durchlässigkeit treten auf für

$$(9) \quad \varphi = 2k\pi.$$

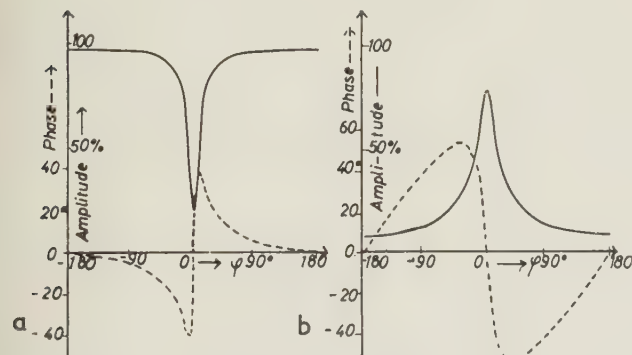


ABB. 5. — Amplitude und Phase von a)  $1r$  und b)  $1d$ .

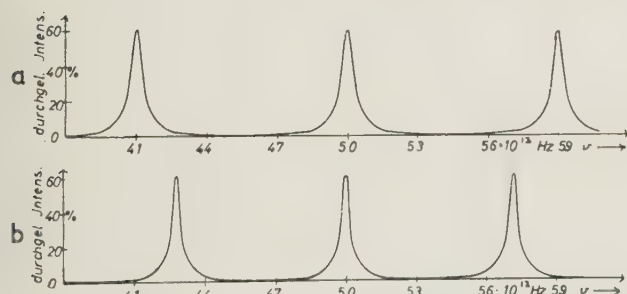


ABB. 6. — Ausschnitte aus den Durchlässigkeitspektren zweier Filter:

- a) Filter I:  $k = 5$ ;  $H_I = 884 \text{ m}\mu$ .  
 Maxima von  $|1d|$  für  $\nu = 5,00$ ;  $5,87$ ;  $6,74$  usw.  
 $\times 10^{14} \text{ sec}^{-1}$ .  
 Halbwertsbreite (bei  $5 \times 10^{14} \text{ sec}^{-1}$ )  $8 \text{ m}\mu$ .  
 Zehntelwertsbreite  $26 \text{ m}\mu$ .
- b) Filter II:  $k = 6$ ;  $H_{II} = 1033 \text{ m}\mu$ .  
 Maxima von  $|1d|$  für  $\nu = 5,00$ ;  $5,73$ ;  $6,46$  usw.  
 $\times 10^{14} \text{ sec}^{-1}$ .  
 Halbwertsbreite (bei  $5 \times 10^{14} \text{ sec}^{-1}$ )  $7 \text{ m}\mu$ .  
 Zehntelwertsbreite  $22 \text{ m}\mu$ .

$k$  heiße "Ordnung" des Maximums. Ist eine auszufilternde Wellenlänge gegeben, so dient (9) zur Bestimmung des Abstandes der Metallschichten.

Die Ordnung  $k$  ist dabei noch frei wählbar und bestimmt nach Gl. (9) die Breite der Linie, da ja  $1r$  und  $1d$  Funktionen von  $\varphi$  allein sind. Erhöht man bei festem  $\lambda$  die Ordnung  $k$ , indem man den Abstand  $H$  der Metallschichten vergrößert, so erhält man schärfere Linien. Diese rücken dabei näher zusammen (sie sind in der  $\nu$ -Skala äquidistant, s. Abb. 6).

Als Beispiele sind zwei Durchlässigkeitsfilter für  $\lambda = 600 \text{ m}\mu$ ,  $\nu = 5 \times 10^{14} \text{ sec}^{-1}$ ,  $k = 5$  und  $k = 6$  berechnet. Abb. 2a und 4-6 stellen die Ergebnisse der Rechnung graphisch dar. Abb. 2a, 4 und 5 entsprechen numerisch dem Filter I von Abb. 6.

4. Das Doppelfilter. — Wir betrachten ein Doppelfilter aus 3 identischen Metallschichten nach Abschnitt 3 in den Abständen  $H_I$  und  $H_{II}$  (Abb. 2b). Wir betrachten es als Durchlässigkeitsfilter, berechnen also nur seine Durchlässigkeit  $2d$  (im Sinne von Gl. (1), bezogen auf  $z_0 = 0$ ).

Um Formel (2) anwenden zu können, fassen wir zwei Schichten zum System I zusammen und fassen die dritte als System II auf (Abb. 2b). Dann ist

$$\begin{aligned} 2d &= \frac{d_I^0 d_{II}^0}{1 - r_I^0 r_{II}^0} \\ &= \frac{1d^0 o d^0}{1 - 1r^0 e^{2\rho_0(H_I-h)} o r^0 e^{-2\rho_0(H_I+H_{II})}} \\ &= \frac{1d^0 o d^0}{1 - 1r^0 o r^0 e^{-2\rho_0(H_{II}-h)}} \\ (10) \quad 2d &= \frac{(od)^3}{N_I N_{II} + K} \end{aligned}$$

mit

$$N_k = \{ 1 - |or|^2 e^{-i\varphi_k} \}; \quad K = |or|^2 |od|^2 e^{-i(\varphi_I + \varphi_{II})}$$

[Bedeutung von  $\varphi_k$  siehe Formel (6)].

Ohne  $K$  im Nenner und bis auf einen Faktor  $(od)^2$  stellt (10) offensichtlich das Produkt der Durchlässigkeiten von zwei Einfachfiltern im Sinne von Abschnitt 3 dar.  $K$  ist also eine Art Kopplungsglied. <sup>1)</sup> In der Gauss'schen Zahlenebene ist  $K$  ein Kreis.

Gibt man den beiden Elementarfiltern zusammenfallende Maxima, so erhält man bei der betreffenden Frequenz auch für das Doppelfilter ein Hauptmaximum, das jedoch durch das Kopplungsglied aus einer scharfen Linie zu einer Bandfilterkurve verformt wird, wie man sie ähnlich in der Hochfrequenztechnik kennt. Abb. 2b zeigt das in der Gauss'schen Zahlenebene und stellt zugleich ein Rechenbeispiel graphisch dar ( $H_I$  und  $H_{II}$  gleich denen von Abb. 6, Hauptmaximum bei  $\lambda = 600 \text{ m}\mu$ ). Die Halbwertsbreite beträgt  $17 \text{ m}\mu$ , die Zehntelwertsbreite  $24 \text{ m}\mu$ . Man hat also bei einer gewissen Verbreiterung eine wesentlich höhere Flankensteilheit im Vergleich zum Einfachfilter.

Störend mag in manchen Fällen das tiefe Minimum

<sup>1)</sup> Das graphische Verfahren wurde von WOLTER in einer Vorlesung WS 1954/55 angegeben [5].

in der Bandmitte sein. Um es zu beseitigen, muss man die geschwungene Kurve in Abb. 2b, d. h. den ersten Nennersummanden von Gl. (10), mehr in die positive

$\left\{ 1 - |o r|^2 e^{-i\varphi} \right\}$  bleibt dann unter  $45^\circ$ , die von  $\left\{ 1 - |o r|^2 e^{-i\varphi} \right\}^2$  unter  $90^\circ$ .

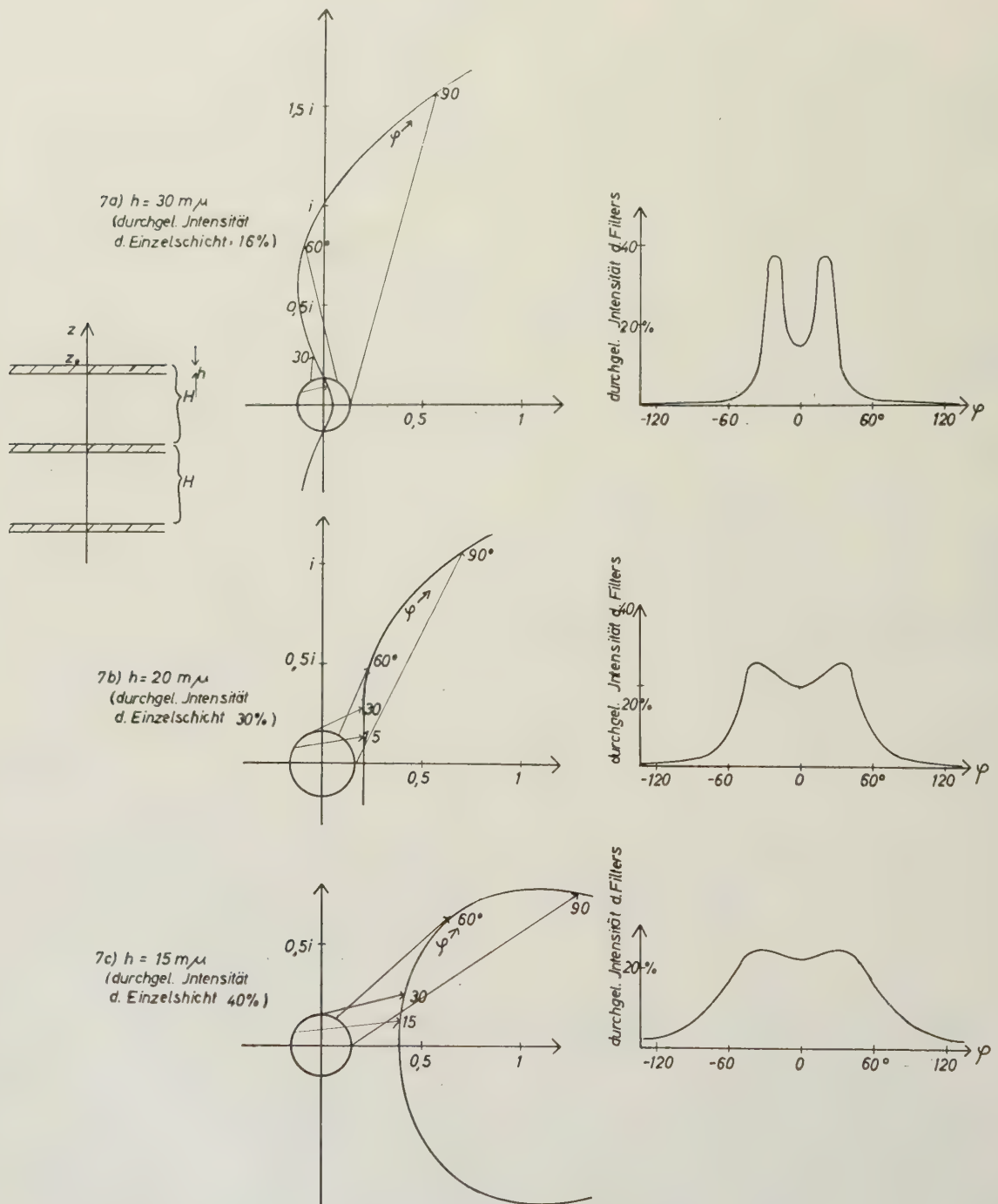


ABB. 7. — Doppelfilter mit verschieden starker Kopplung. Erste Spalte: Aufbau des Filters. Zweite Spalte: Der Nenner von  $2D^*$  (s. Gl. (11)) in der Zahlenebene. Dritte Spalte: Durchgelassene Intensitäten in % der eingefallenen.

Halbebene ziehen und die Einfeldung bei  $\varphi = 0$  ein-ebnen. Das gelingt aber offensichtlich für kleinere  $|o r|$ , wie man aus Abb. 2a ablesen kann: Die Phase von

Die Verhältnisse sind in Abb. 7 dargestellt.  
Für a)  $h = 30 \text{ m}\mu$ ,  $n_m = 0,12 - 4,15 i$ ,



$$b) h = 20 \text{ m}\mu, n_m = 0,14 - 4,20 \text{ i.}$$

$$c) h = 15 \text{ m}\mu, n_m = 0,14 - 4,25 \text{ i.}$$

(Werte aus MAYER l. c.) ist

$$(11) \quad {}_3\delta^* = \frac{(\delta\delta)^3}{\left\{ 1 - |o r|^2 e^{-i\varphi} \right\}^2 + |o r|^2 |\delta\delta|^2 e^{-2i\varphi}}$$

nach gerechneten Werten als Funktion von  $\varphi$  dargestellt. Zur Vereinfachung wurden also die beiden Elementarfilter gleich gemacht ( $\varphi_I = \varphi_{II} = \varphi$ ).

Abb. 7 zeigt, wie die Kurven sich für dünnere Metallschichten ins Positive ziehen; daneben gibt sie den Intensitätsverlauf wider.

Die Eindellung ist für  $h = 20 \text{ m}\mu$  nur noch ganz gering. Die Einebnung wird durch Linienverbreiterung und geringere Flankensteilheit erkauft.

**5. Das Dreifachfilter.** — Endlich sei ein Dreifachfilter besprochen. Es besteht aus 4 Metallschichten im Sinne von Abschnitt 3 in den Abständen  $H_I$ ,  $H_{II}$  und  $H_{III}$  (Abb. 2c). Wir fragen wieder nur nach der Durchlässigkeit  ${}_3\delta$ , gehen dazu auf Gl. (2) zurück und fassen je zwei Metallschichten zu den Systemen I und II zusammen, deren Durchlässigkeit Gl. (7) angibt. Dann ist die Gesamtdurchlässigkeit

$${}_3\delta = \frac{\delta_I^o \delta_{II}^o}{1 - r_I^{o'} r_{II}^o}$$

$$(12) \quad {}_3\delta = \frac{(\delta\delta)_I (\delta\delta)_{II}}{1 - (r_I) (r_{II}) e^{-2\rho_0(H_{II}-h)}}$$

Eine elementare Zwischenrechnung ergibt

$$(13) \quad {}_3\delta = \frac{(\delta\delta)^4}{N_I N_{II} N_{III} + K^*}$$

mit

$$N_k = \left\{ 1 - |o r|^2 e^{-i\varphi_k} \right\}; \quad K^* = |o r|^2 |\delta\delta|^2 e^{-i\varphi_{III}} \left\{ e^{-i\varphi_I} + e^{-i\varphi_{II}} - [|\delta\delta|^2 + 2|o r|^2] e^{-i(\varphi_I + \varphi_{II})} \right\}.$$

Auch hier kann man in  $K^*$  das Kopplungsglied erkennen. Der ganze Nenner ist aber so kompliziert gebaut, dass auf seine Darstellung in der Zahlenebene verzichtet sei. Zur numerischen Berechnung eines Beispiels geht man besser von (12) aus. Auf diesem Wege ergab sich mit  $H_{III} = H_I$ ,  $H_{II}$  von Abb. 6, die Kurve der Abb. 2c.

Auch für das Dreifachfilter ergibt sich, wie zu erwarten, eine Bandfilterkurve mit hoher Flankensteilheit, das Band erscheint jedoch gegenüber dem Doppelfilter schon etwas geglättet. Im sichtbaren Teil des Spektrums ergab eine graphische Abschätzung Nebenmaxima mit bis zu 17% der Hauptmaximumsintensität.

Durch Variation von Dicke und Abstand der Metallschichten, durch Wahl anderer Metalle und Dielektrika, endlich durch weitere Erhöhung der Schichtzahl, können die Filterkurven natürlich weitgehend abgeändert werden. So kann man die Linien etwa verschärfen, indem man die Ordnung der Elementarfilter erhöht. Wie man durch Änderung der Metallschichtdicken die Linienstruktur ändern kann, wurde bereits in Abschnitt 4 erörtert.

Herrn Prof. Dr. WOLTER danke ich für die Anregung zu der Arbeit und für anregende Diskussionen.

#### LITERATUR

- [1] W. GEFFCKEN, *DRP.* Nr 716 153 (1942).
- [2] K. SCHUSTER, *Ann. d. Phys.* (6), **4**, 1949, 352.
- [3] W. GEFFCKEN, *Zeitschr. f. ang. Phys.*, **6**, 1954, 249.
- [4] H. MAYER, *Physik dünner Schichten*, Stuttgart, 1950.
- [5] H. WOLTER, *Hdb. Physik*, Bd 24 (1956).

*Manuscrit reçu le 15 août 1955.*

## Distribution des unités sensorielles dans l'extrafovéa \*

J. OLIVA et M. AGUILAR

Instituto de Optica, Madrid

SOMMAIRE. — Nous étudions ici la symétrie circulaire de la rétine extrafovéale (jusqu'à 28° de la fovéa) en ce qui concerne l'acuité visuelle étroitement liée à la distribution des unités sensorielles dans la rétine.

Les résultats obtenus nous montrent que cette symétrie existe et qu'il y a augmentation de la grandeur des unités sensorielles et diminution de l'acuité visuelle à mesure que nous nous éloignons de la fovéa.

ZUSAMMENFASSUNG. — Die Arbeit prüft die Sehschärfe und die damit eng verbundene Verteilung der Wahrnehmungselemente auf der Netzhaut hinsichtlich ihrer symmetrischen Verteilung auf dem extrafovealen Teil der Netzhaut bis zu einem Abstand von 28° von der Fovea. Die Ergebnisse zeigen, dass eine solche Kreissymmetrie besteht und dass je mehr man sich von der Fovea entfernt, die Grösse der Wahrnehmungselemente zunimmt und die Sehschärfe entsprechend verringert wird.

SUMMARY. — A study is made of the circular symmetry of visual acuity of the extra-foveal retina (up to 28° from the fovea), which is closely allied to the distribution of sensory units in the retina.

The results obtained show this symmetry exists, and that there is an increase in the size of sensory units and a reduction of visual acuity as the distance from the fovea increases.

La distribution des cônes et des bâtonnets dans l'extrafovéa a déjà été étudiée objectivement par OSTERBERG qui trouve une distribution asymétrique. Mais ce chercheur ne parle pas des unités sensorielles qui nous intéressent particulièrement dans cette étude.

De nombreux travaux ont été faits sur l'acuité visuelle et les seuils dans l'extrafovéa. Mais presque tous se limitent au méridien horizontal et quelques-uns seulement au méridien vertical ; c'est pourquoi nous avons pensé intéressant de réaliser une étude sur une large zone de l'extrafovéa.

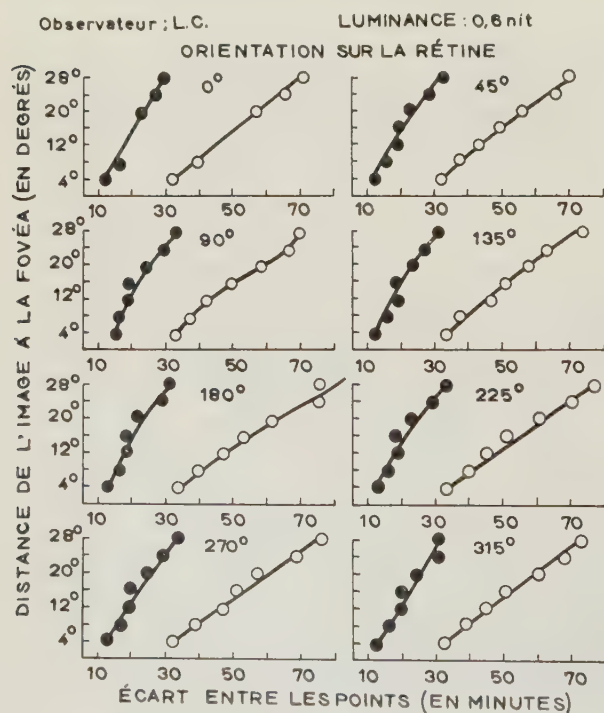


FIG. 1.

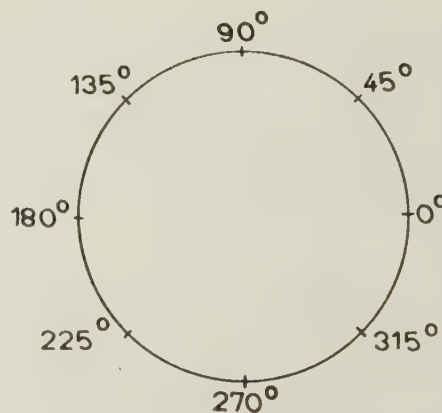


FIG. 2.

La technique utilisée fut la suivante.

L'observateur qui regarde deux points lumineux sur fond sombre, situés sur une même horizontale et se rapprochant l'un de l'autre doit signaler le moment où il ne distingue plus les deux points, mais où il aperçoit un trait horizontal, et ensuite, le moment où ce trait lumineux se transforme en un point. Toutes ces mesures se répètent avec des extrafovéalités différentes ; les résultats obtenus, pour un observateur et une luminosité donnés, sont portés sur les courbes de la figure 1.

Sur chacun de ces diagrammes on porte, en ordonnées les extrafovéalités exprimées en grades et, en abscisses, l'écart des 2 points exprimé en minutes. L'orientation de la vision est indiquée sur chaque diagramme. Elle est mesurée à partir de la vision nasale à laquelle correspond l'angle 0°. La figure 2 fera aisément comprendre l'orientation à laquelle correspond chacune des courbes. Au centre de la circonférence, sont placés les deux points constituant le test des expériences, et sur la circonférence se déplace la lumière rouge de fixation. L'observateur regarde avec l'œil gauche.

\* Ce papier a été présenté partiellement à l'occasion de la Réunion Internationale d'Optique de Florence (septembre 1954).



Les courbes de droite nous donnent la distance à laquelle nous cessons de percevoir le trait pour voir seulement un point. Ces courbes sont en relation étroite avec la grandeur des unités sensorielles, car nous croyons que le fait de voir un seul point peut faire supposer que les deux images tombent sur la même unité sensorielle. Pour confirmer qualitativement ce fait nous avons réalisé quelques expériences en donnant aux points un écart inférieur à celui qui permet de voir un trait et en présentant ces points alternativement à l'observateur ; celui-ci a vu un point constamment immobile au lieu d'un point vibrant, ce qui, selon nous, devrait avoir lieu si les images des points tombent sur des unités sensorielles contiguës. En effet, lorsque nous avons fait la même expérience en donnant aux points un écart pour lequel ceux-ci apparaissent comme un trait l'observateur a noté une vibration d'amplitude croissante avec l'écart des points.

Les dimensions réelles de ces unités ne coïncident pas exactement avec les dimensions trouvées expérimentalement, et cela pour les raisons suivantes :

1) Les mouvements de l'image rétinienne pendant la fixation de l'œil, pour une durée de vision de 1/3 de seconde (celle employée dans nos expériences) peuvent faire varier les résultats d'environ deux minutes au maximum (selon RIGGS [2]).

2) Les images des points observés, quoique faisant abstraction de la diffraction, ne sont pas ponctuelles, mais au contraire, relativement étendues ainsi que l'a signalé Y. LE GRAND [3].

Les dimensions des unités sensorielles que nous avons trouvées coïncident avec celles obtenues dans un travail antérieur [4] employant une technique toute différente.

La forme de ces courbes confirme une fois de plus l'augmentation de la grandeur des unités lorsque grandit l'extrafovéalité.

Quoique apparemment ces résultats ne coïncident pas avec ceux de CLEMMENSEN (1944) [5], selon lesquels, les unités sensorielles sont pratiquement égales à partir de 3° de la fovéa, en réalité il n'y a pas contradiction, bien que les unités sensorielles auxquelles nous nous référons puissent être celles que nous avons appelées unités sensorielles secondaires dans des travaux antérieurs [4] ; dans chaque cas notre interprétation coïncide avec celle de CLEMMENSEN lorsqu'il dit qu'à partir de 3° la sommation se vérifie grâce aux chemins transversaux de la rétine, permettant l'interconnection d'un nombre supérieur de récepteurs.

Les courbes de droite (fig. 1) nous donnent la valeur inverse de l'acuité visuelle et leur allure nous indique la variation de l'acuité visuelle avec l'extrafovéalité.

Sur la figure 3, les ordonnées indiquent la séparation des points en minutes, les abscisses, les différentes orientations de la vision et dans la partie supérieure on exprime en degrés l'angle d'extrafovéalité.

La courbe inférieure correspond au passage de la vision du point à celle du trait, et la courbe supérieure au passage du trait à deux points séparés.

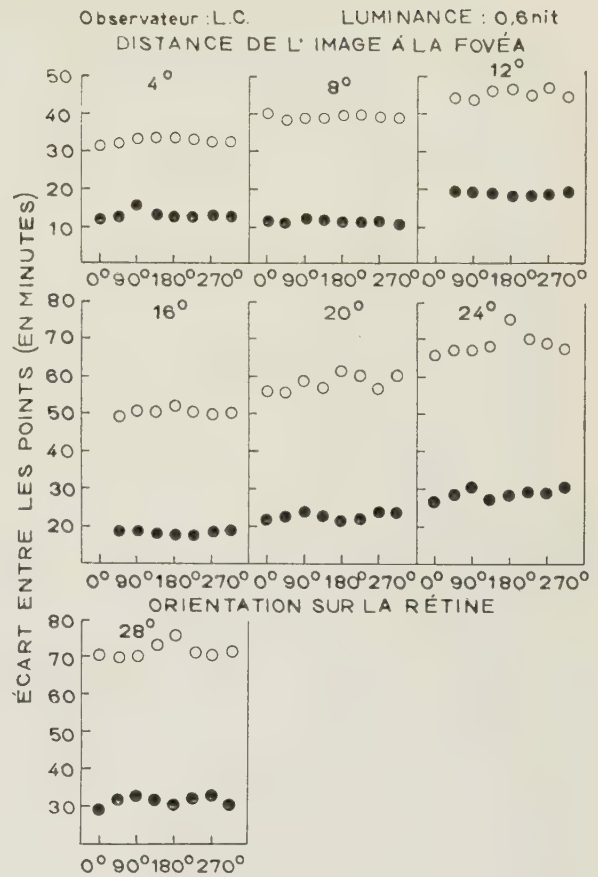


FIG. 3.

Ces courbes montrent clairement la symétrie circulaire existant autour de la fovéa, au moins jusqu'à 28°, pour ce qui concerne l'acuité visuelle. Cependant en ce qui concerne la symétrie nous avons seulement expérimenté jusqu'à 28°, nous avons désiré déterminer pour deux des observateurs et avec une seule orientation (vision nasale) la prolongation des courbes de la figure 1, jusqu'à 60°. Les résultats correspondants sont donnés sur la figure 4, et l'on peut voir que le comportement des unités sensorielles et de l'acuité visuelle est analogue à celui qui pourrait être prévu par extrapolation des courbes déjà obtenues.

Dans tout ce qui précède nous nous référons à des expériences où les points lumineux étaient situés sur la même horizontale et par suite lorsque nous parlons de la variation de grandeur, nous nous référons à la dimension horizontale de l'unité sensorielle. Pour vérifier s'il y avait des variations verticalement, nous avons répété quelques-unes des expériences avec les points situés alternativement en position horizontale et verticale, et nous avons constaté que l'écart des points nécessaire dans les deux positions était le même. Ceci nous fait supposer une forme à peu près circulaire de ces unités. Bien que ces résultats semblent contredire ce que l'on sait de l'astigmatisme nocturne (6) qui est en général fonction de l'orienta-

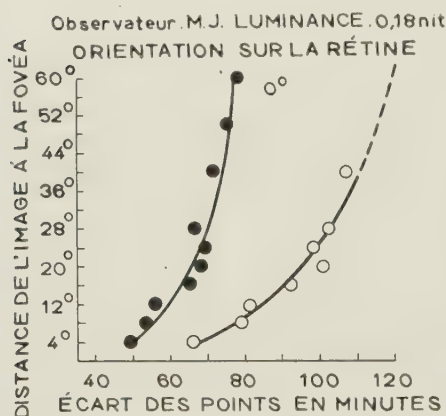
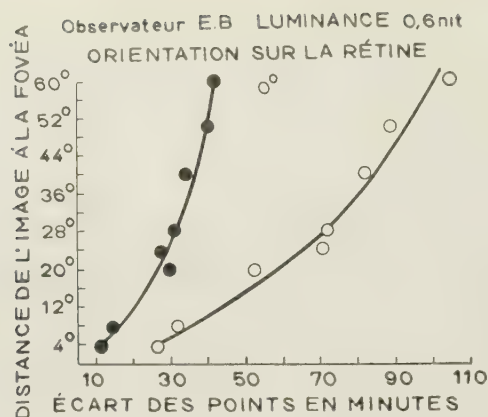


FIG. 4.

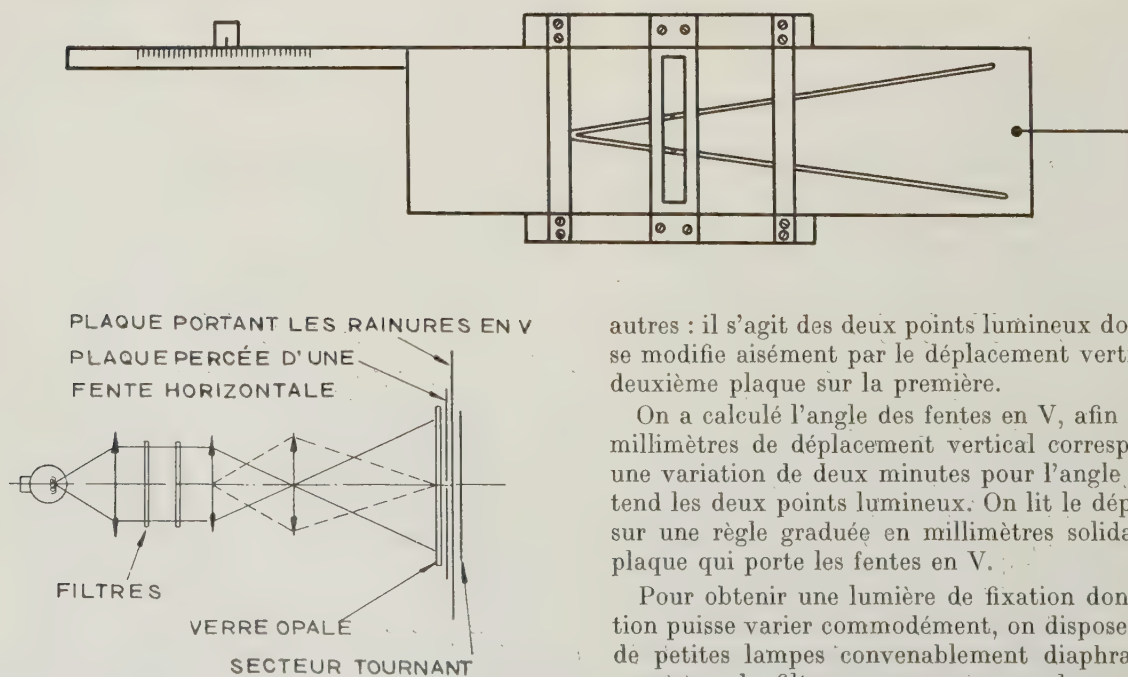


FIG. 5.

tion de la lumière de fixation, cette contradiction est seulement apparente, parce que le phénomène d'astigmatisme se manifeste dans des expériences où intervient la sommation et ce n'est pas le cas ; dans ces expériences ce qui influe est la grandeur des unités sensorielles, ainsi que la densité de distribution de ces unités.

En résumé nous pouvons dire que

- 1) les unités sensorielles dans l'extrafovée ont une forme approximativement circulaire ;
- 2) la grandeur de ces unités varie en fonction de l'extrafovéalité dans toute la zone étudiée (jusqu'à 60°) ;
- 3) cette variation de grandeur garde une symétrie circulaire autour de la fovée ;
- 4) la densité des unités sensorielles diminue lorsque l'extrafovéalité augmente, et cette densité se répartit aussi symétriquement par rapport à la fovée.

**Dispositif employé.** — La manière dont s'obtiennent les deux sources est indiquée sur la figure 5. Un verre opale de grande homogénéité est éclairé par un système d'illumination de KÖHLER qui permet d'intercaler sur le parcours parallèle des rayons, les filtres nécessaires pour obtenir l'intensité désirée. On place devant le verre opale, une plaque munie d'une fente horizontale mesurant 1 mm de largeur. Cette plaque est pourvue de plusieurs coulisses lui permettant le glissement sur une deuxième plaque munie de deux fentes de même largeur que la précédente et en forme de V, de sorte que l'ensemble des deux plaques ne laisse passer la lumière que par les deux points d'intersection de la fente horizontale avec les deux

autres : il s'agit des deux points lumineux dont l'écart se modifie aisément par le déplacement vertical de la deuxième plaque sur la première.

On a calculé l'angle des fentes en V, afin que deux millimètres de déplacement vertical correspondent à une variation de deux minutes pour l'angle qui sous-tend les deux points lumineux. On lit le déplacement sur une règle graduée en millimètres solidaire de la plaque qui porte les fentes en V.

Pour obtenir une lumière de fixation dont la position puisse varier commodément, on dispose une série de petites lampes convenablement diaphragmées et pourvues de filtres rouges, sur une baguette rigide mais légère pouvant tourner autour d'un axe perpen-



diculaire représenté par la ligne horizontale joignant l'œil de l'observateur au point central des deux lumières.

Les lampes rouges sont placées de telle sorte que le passage de l'une à la suivante représente une variation de l'angle de fixation de 2°. On peut arriver à 30° de la fovéa. Un index permet d'orienter la baguette dans des directions formant entre elles 45°.

Une résistance variable permet d'établir une intensité lumineuse convenable.

L'observateur est placé à 2,60 m des points lumineux, son menton repose sur un appareil dont l'orientation peut varier, afin que sa face soit toujours dirigée vers la lumière rouge.

**Mode opératoire.** — On choisit premièrement une intensité des points qui ne soit pas trop forte et ne produise pas une sensation d'éblouissement à l'observateur en vision extrafovéale et adapté à l'obscurité depuis 40 minutes. L'entraînement des observateurs fut laborieux pendant les premiers jours et même les premières semaines, la précision obtenue fut faible.

On a amélioré la précision en éclairant les points lumineux par intermittence : un secteur rotatoire est placé immédiatement devant les points qui apparaissent et disparaissent alors simultanément. Le temps d'éclairement fut de 1/3 de seconde, avec un intervalle de 1,7 seconde.

A part les 40 minutes d'adaptation à l'obscurité, les séances ont duré approximativement une heure. Chaque jour, au commencement, on a situé la lumière rouge dans les positions 16° et 4° de la fovéa sur le côté temporal, de sorte que ces deux lectures ont rendu possible la comparaison entre les différents jours dont quelques-uns furent exclus par suite de différences notables.

L'intensité des points lumineux a été de  $6 \times 10^{-7}$  candelas dans la première série d'expériences et de  $1,8 \times 10^{-7}$  candelas dans la seconde.

On a obtenu chaque point par la moyenne de quatre lectures réalisées en quatre jours.

Pour réaliser ces expériences, on a varié le plus possible l'ordre d'exposition afin d'éviter que la fatigue puisse influencer les résultats finaux.

Nous avons parlé d'une confirmation qualitative de ce que lorsque les deux points apparaissent comme n'en formant qu'un seul, leurs images se forment sur une même unité sensorielle.

Nous avons réalisé cette expérience de la façon suivante. Devant les points et très près de ceux-ci

on a placé un disque giratoire cachant alternativement un point puis l'autre, de sorte que l'apparition de l'un d'eux correspondait à la disparition de l'autre. L'observateur regardait extrafovéalement et auparavant on plaçait les points suffisamment près pour que l'on en perçoive seulement un. Pendant la rotation du disque on ne notait aucune vibration du point lumineux. Au contraire, lorsque la séparation entre les points était telle que l'observateur voyait un trait lumineux, pendant la rotation du disque, on notait une vibration du point. L'amplitude de cette vibration devenait plus grande si la séparation des points augmentait.

Nous avons également fait des expériences pour voir quelles modifications se seraient introduites dans les valeurs des séparations angulaires entre les points, si l'on plaçait ceux-ci sur une verticale au lieu d'une horizontale.

Pour cela nous avons monté un dispositif où l'on obtenait les points pour un procédé analogue mais en substituant les deux fentes lumineuses, en V par deux séries de petits orifices d'égale grandeur, disposés également en V. Cette plaque rectangulaire pouvait glisser sur une autre qui portait une fente coupant les deux rameaux du V afin qu'en chaque position apparaissent seulement deux points lumineux. Le système était tournant et l'on passait ainsi facilement de la position horizontale à la position verticale.

Nous avons opéré en vision extrafovéale, temporale et frontale de 8° et 20°. Nous avons vu que ces valeurs obtenues pour la séparation des points étaient indépendantes de la position horizontale ou verticale de ceux-ci, aussi bien quand il s'agissait de voir la séparation maximum pour que les points se voient comme un seul, que lorsqu'on déterminait la séparation maximum des points qui permet de les voir confondus en un trait lumineux.

#### BIBLIOGRAPHIE

- [1] ØSTERBERG, *Acta Ophthalm.*, supp., **6** (1935).
- [2] JIGGS, ARMINGTON and RATLIFF, *J. O. S. A.*, **44**, 1954, 315.
- [3] LE GRAND, Y., *Rev. d'Opt.*, 1937, 201.
- [4] M. AGUILAR et J. YUNTA, *An. Real. Soc. Fis. Quim.*, 1945.
- [5] CLEMENSEN, 1944, cité par HARTRIDGE dans « *Recent advances in the Physiology of vision* » p. 19, 1950.
- (6) J. M. OTERO, M. AGUILAR et J. YUNTA, *An. Real. Soc. Fis. Quim.*, **47**, 1951, 295.

Manuscrit reçu le 6 décembre 1955

## Visual experiments with a vibrating test object (\*)

ANNA MARIA ERCOLES. ADRIANA FIORENTINI and G. TORALDO DI FRANCA  
Istituto Nazionale di Ottica-Arcetri-Firenze

**SUMMARY.** — *In order to study the influence of eye movements on vision, detection and resolution have been measured with a vibrating test object. The experimental data obtained by three observers have been plotted both against frequency and amplitude of the vibration. It seems possible to conclude that eye movements, even of a very high frequency, have a rôle in affecting vision of moving objects.*

**SOMMAIRE.** — *Afin d'étudier l'influence des mouvements de l'œil sur la vision, on mesure la limite de perception et le pouvoir de résolution avec un test objet vibrant. Les données expérimentales obtenues par trois observateurs ont été représentées en fonction de la fréquence et de l'amplitude de la vibration.*

*Il semble permis de conclure que les mouvements des yeux, même de très grandes fréquences, jouent un rôle dans la vision d'objets en mouvement.*

**ZUSAMMENFASSUNG.** — *Um den Einfluss der Augenbewegungen auf das Sehen zu untersuchen, wurden die Erkennbarkeit und die Auflösung an schwingenden Testobjekten gemessen. Die Versuchsergebnisse, die man an drei Beobachtern erhalten hat, wurden beide aufgetragen gegen die Amplituden und Frequenz der Schwingung. Man darf wohl schliessen, dass die Augenbewegungen selbst bei sehr hohen Frequenzen für das Sehen bewegter Objekte eine Rolle spielen.*

**1. Introduction.** — A question at present of great interest in the field of physiological optics is the study of involuntary movements of the eye. A number of researchers have recently succeeded in measuring eye movements, by different methods [1], [2], [3], [4], [5], [6], [7].

There is some evidence that eye movements have an influence on vision. Many interesting researches on this subject can be mentioned [8], [9], [10]; they prove that the visibility of very fine objects and the resolving power of the eye are greatly reduced when the effect of eye movements is eliminated by maintaining the image of the test object fixed on the retina.

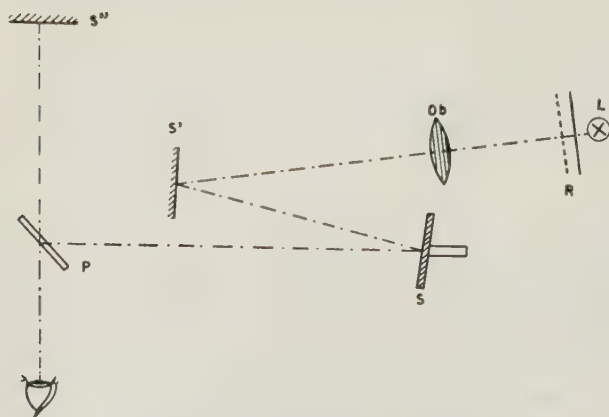


FIG. 1. — Experimental arrangement.

Some information regarding the influence of both voluntary and involuntary eye movements on vision can be obtained by another method. Visual acuity can be measured with a moving test object. The test

object may be a grating, with black and white parallel lines, which is made to oscillate in the direction perpendicular to the lines. Frequency and amplitude of

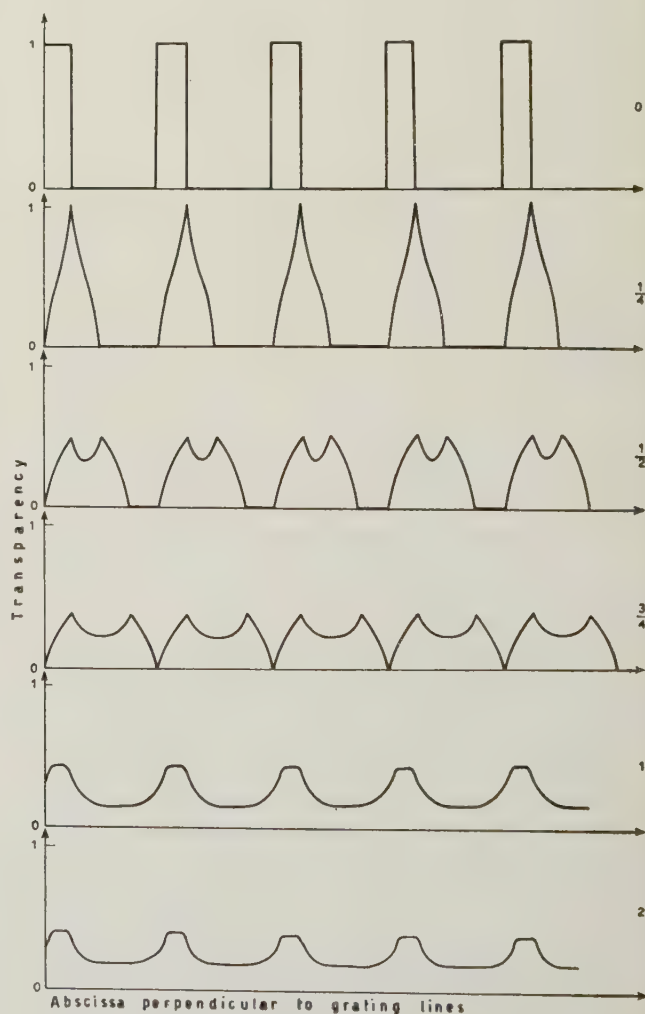


FIG. 2. — Equivalent photometric profiles of grating for different amplitudes of oscillation.

(\*) The research reported in this paper has been made possible through the support and sponsorship of the Air Research and Development Command, United States Air Force, through its European Office under contract No. AF 61 (514)-634 C.

Some preliminary informations about this research have been given at the Florence Meeting on « Problems in Contemporary Optics », 10-15 september 1954.

the oscillation is varied in the same range as those of natural eye movements.

The oscillation of the test object produces a displacement of the image on the retina which combines with the displacement due to the eye movements. Of course, the type of movements which are easiest to realize are regular and periodical, whereas eye movements are more or less irregular. It is impossible to predict the resultant movement. It is possible however, to ascertain by an experimental research whether or not visibility of the test object is affected by the movement; and possibly what values of amplitude and frequency give rise to the greatest effect. This investigation is the subject of the present research.

**2. Experimental method.** — In order to measure visual acuity with a vibrating test object, the following apparatus has been used.

The photographic grating R (fig. 1) with parallel and alternately opaque and transparent lines, is illuminated through an opal glass by the lamp L, fed with d. c. An image of R is focused by the lens Ob very close to the plane mirror S. The observer looks at this image, reflected by a second mirror S'', through the semi-transparent plate P.

Mirror S turns around an axis which makes a small angle with the normal to the mirror. During rotation of S, the image of R undergoes a circular motion, with the same frequency as S. Frequency of rotation is

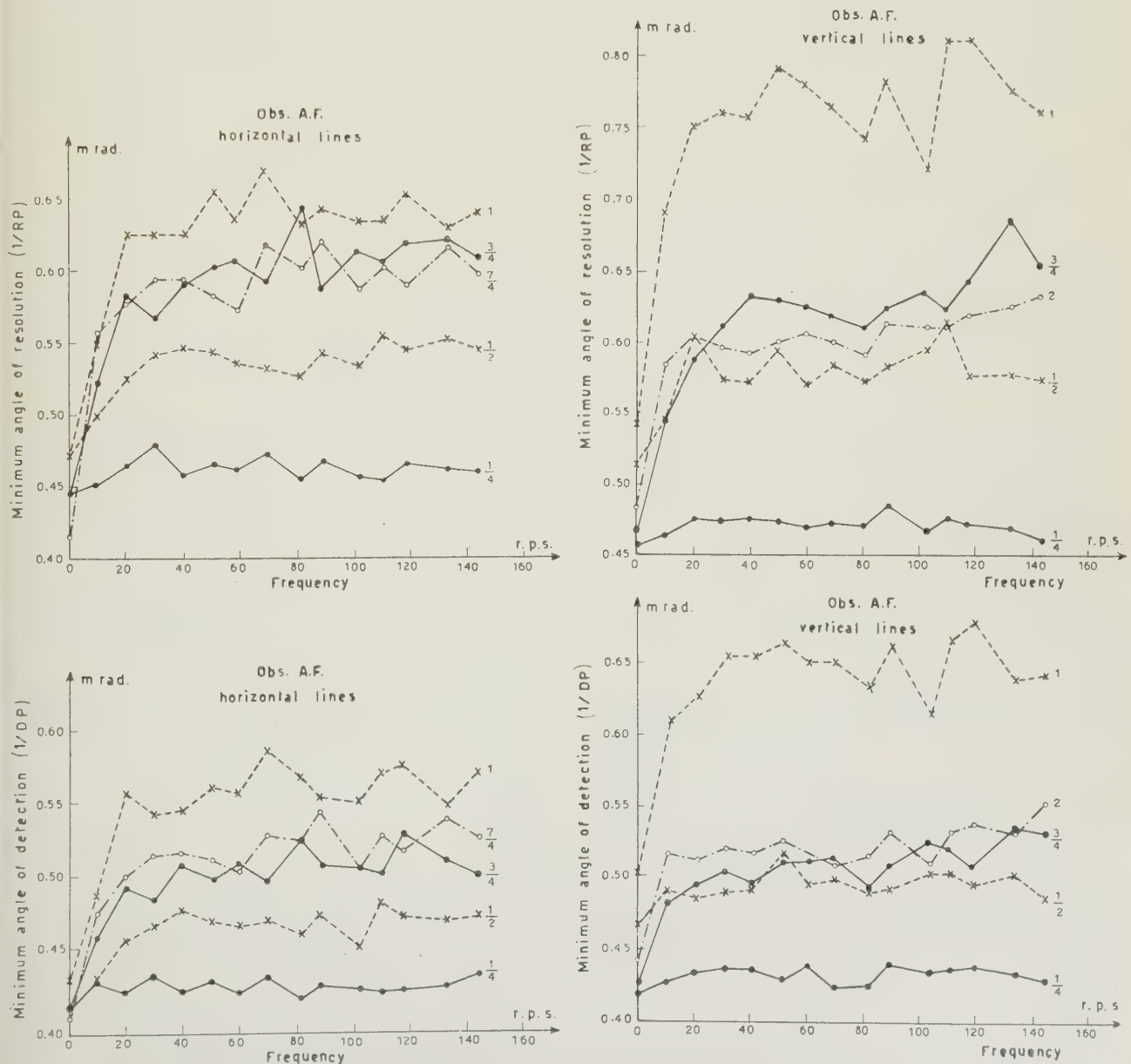


FIG. 3. — Curves of detection and resolution as functions of oscillation frequency for (a) horizontal and (b) vertical grating. Observer A. F.



measured and held constant during each observation by means of a stroboscope. The angular amplitude of the movement can be adjusted by varying either the inclination of S or the distance between S and the image of R (by means of small displacements of mirror S').

The component of the circular motion of the image parallel to the lines of R has obviously no influence on resolution. We must only be concerned with the oscillation perpendicular to the grating lines.

If the frequency of the movement were infinite, the vibrating test object would be equivalent to a fixed test object, with smaller contrast. The luminance curve of the equivalent test object can be worked out, for any given amplitude of the movement. The computations have been performed for the particular case used in our experiment where the opaque stripes of the test object are three times wider than the transparent ones. Their photometric profile has been assumed to be rectangular.

Figure 2 shows the luminance distribution of the motionless test object and of the equivalent test objects corresponding respectively to amplitudes of  $1/4$ ,  $1/2$ ,  $3/4$ , 1 and 2 times the grating spacing  $p$ . The relative luminance is plotted against an abscissa on the grating taken perpendicular to the lines.

If now eye movements are taken into account, one may anticipate that visual acuity when measured with a vibrating test object would not be the same as when measured with a test object at rest having the corresponding luminance distribution. Indeed it is possible as said above that eye movements may combine with the test object movement to alter visibility.

The observations have been performed in the following manner. An operator adjusts the parallel line test object in a vertical or horizontal position. Then he slowly raises the illumination of the grating from zero to a fixed value. (Approximate luminance of

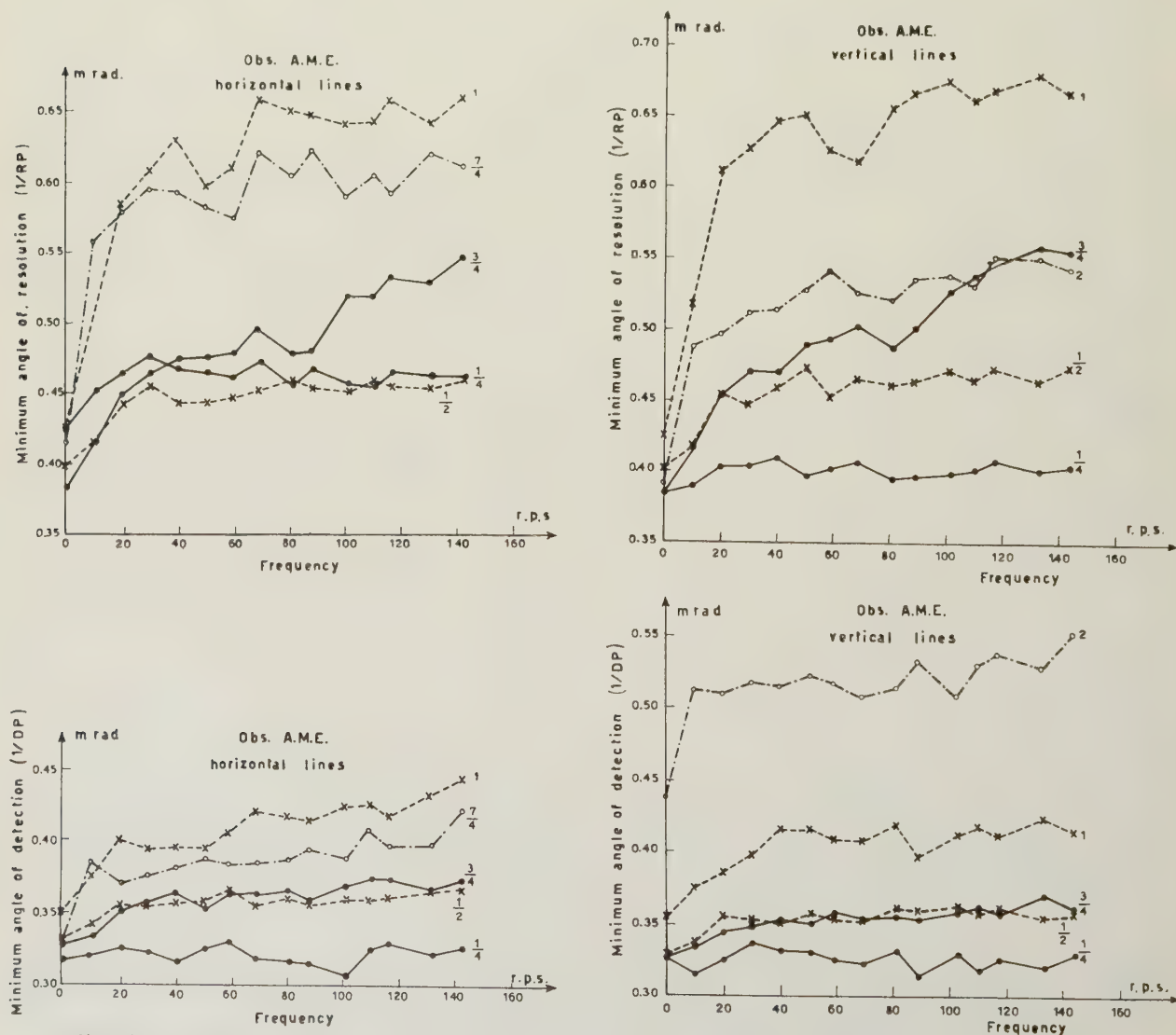


FIG. 1. - Curves of detection and resolution as functions of oscillation frequency for (a) horizontal and (b) vertical grating. Observer A. M. E.

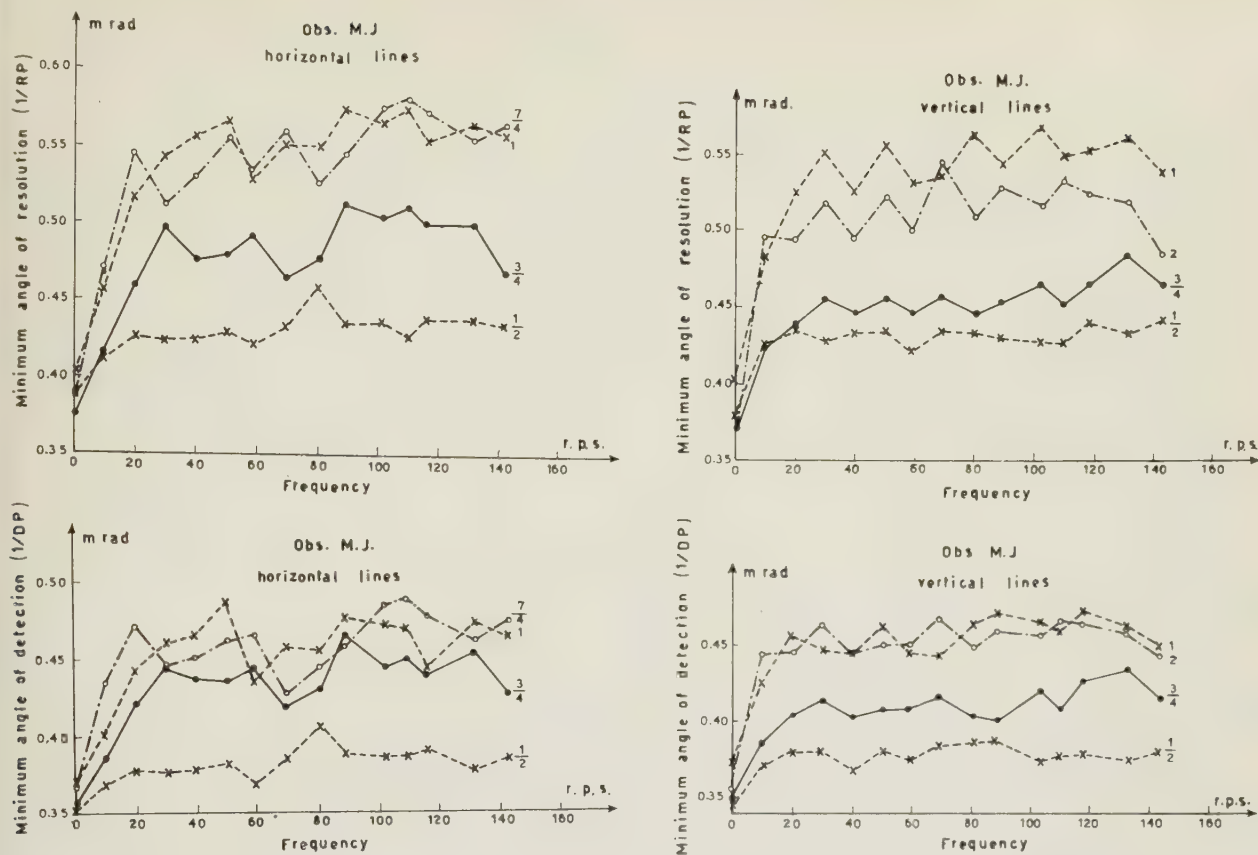


FIG. 5. — Curves of detection and resolution as functions of oscillation frequency for (a) horizontal and (b) vertical grating. Observer M. J.

dark and bright lines 10 nit and 420 nit). The purpose of this procedure is to avoid the effect of an abrupt formation of the image on the retina. The observer, after a fifteen minute dark adaptation, pulls mirror  $S''$  towards himself until he is able to detect the orientation of the lines. The inverse of the angle subtending the ruling spacing  $p$ , for this position of  $S''$  has been termed the « detecting power » (DP).

Next, by pulling  $S''$  still closer, the observer can determine a new position where he is able to resolve the single lines of the grating. The inverse of the corresponding visual angle is the resolving power RP.

The two measurements of DP and RP have been performed, for a given amplitude of the movement, with each of the following set of vibration frequencies: 0, 10, 20, 30, 40, 51, 59, 69, 82, 88, 103, 110, 117, 133, 144 r. p. s.

The succession of different frequencies followed a random order, which was unknown to the observer. Each set of measurements consisted of two observations for each of the above frequencies. Ten sets of observations have been made for each of the five values of the movement amplitude, namely  $1/4$ ,  $1/2$ ,  $3/4$ , 1 and 2 times the grating spacing. Each set of measurements took place in a different day.

3. **Results.** — Figures 3, 4, 5 show the curves obtained

by three observers (A. F., A. M. E., M. J.). The inverses of both DP and RP (visual angle subtended by the grating spacing, in milliradians), are represented in the ordinates. The abscissae represent the frequency of the oscillation in r. p. s. Each point is the mean of 10 independent observations, whose standard deviation is of the order  $\pm 0.01$  milliradian.

Each curve corresponds to a constant amplitude. Amplitudes  $a$  for the horizontal test object are respectively  $1/4$ ,  $1/2$ ,  $3/4$ , 1,  $7/4$  times the spacing  $p$  of the grating. For the vertical test object  $a$  has the values  $1/1/4$ , 2,  $3/4$ , 1, 2 times  $p$ . Observer M. J. has not performed measurements with an amplitude of  $1/4$  the spacing.

From these curves it is an easy matter to obtain other graphs showing DP and RP as functions of the amplitude at constant frequency (fig. 6, 7, 8). As it is advisable to exclude from our data the accidental variations of visual acuity due to a change of the physiological state of the observer (see the next paragraph) we have represented on the vertical axis the difference between  $1/DP$  and  $1/RP$  as measured with the moving test object and with the fixed one.

4. **Discussion of the results.** — When trying to interpret the results one should take into account that, due to the duration of the observations, the

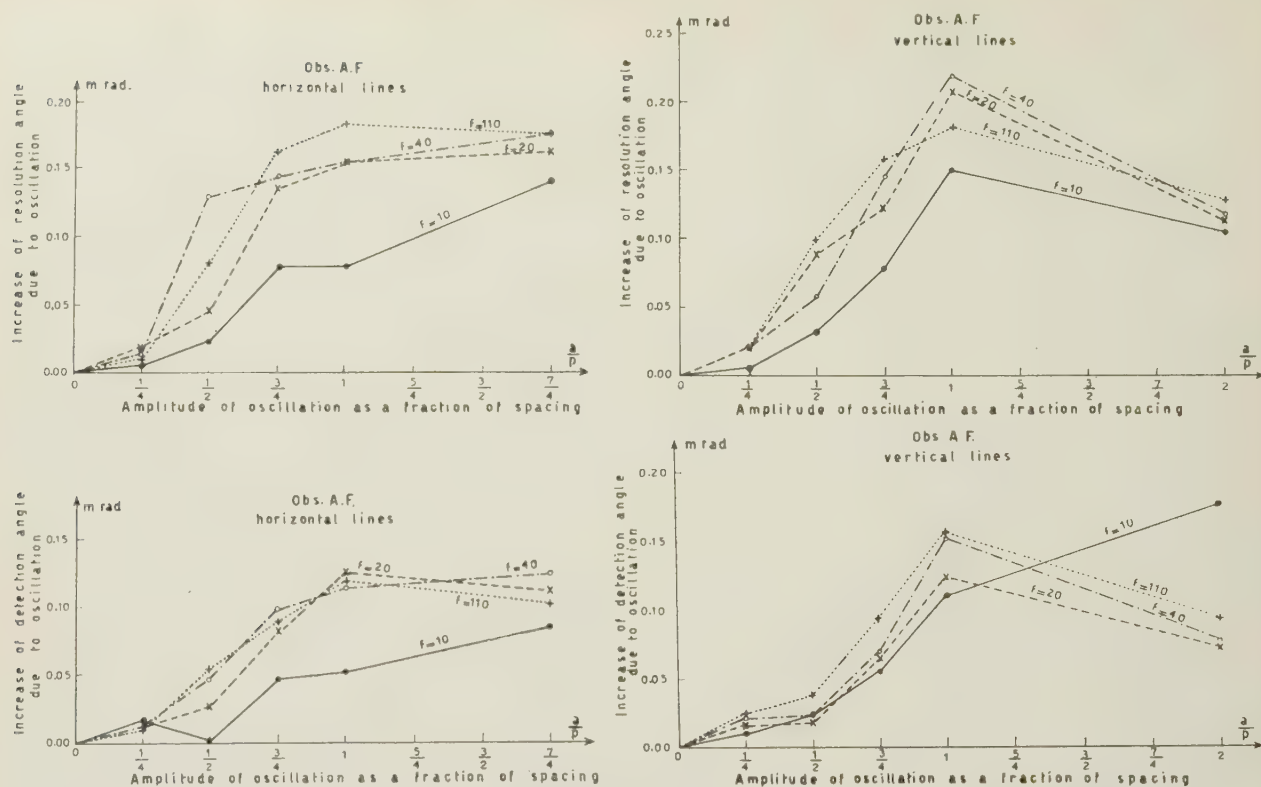


Fig. 6. — Increase of detection and resolution angles due to oscillation, as functions of amplitude, for (a) horizontal and (b) vertical grating. Observer A. F.

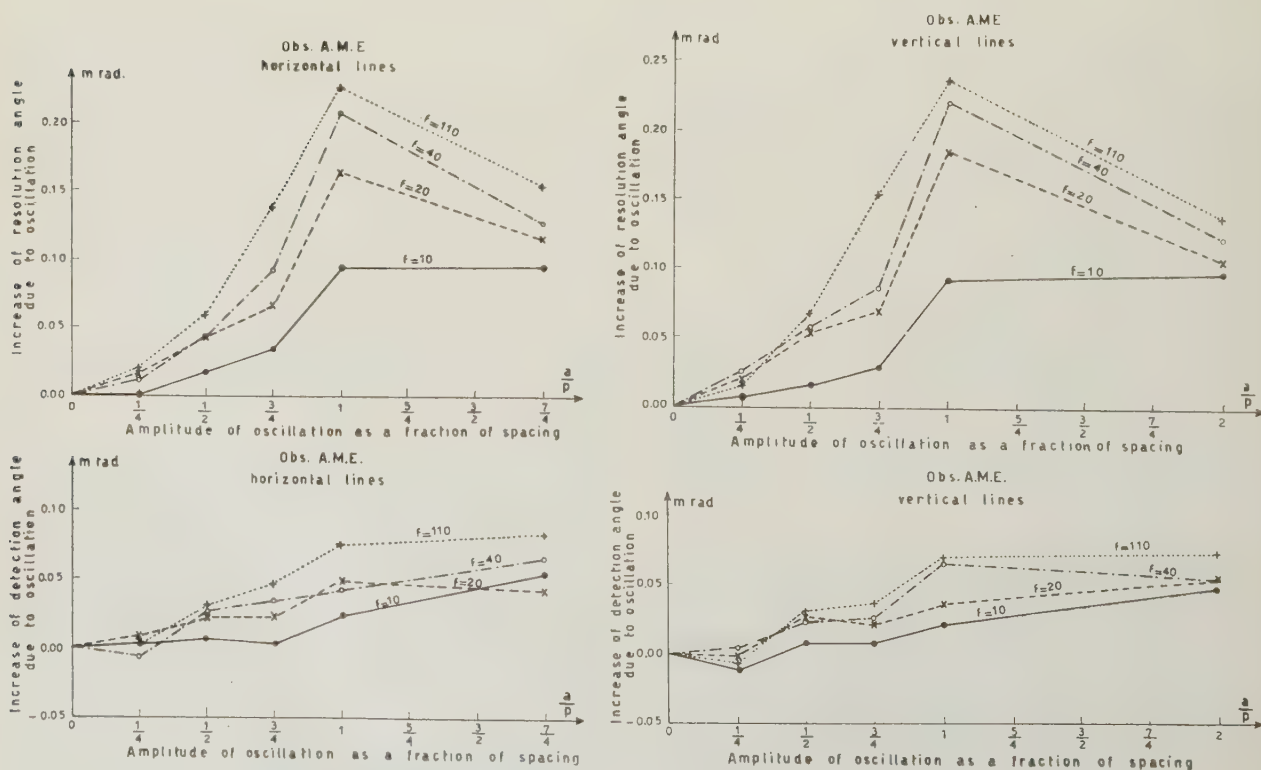


Fig. 7. — Increase of detection and resolution angles due to oscillation, as functions of amplitude, for (a) horizontal and (b) vertical grating. Observer A. M. E.



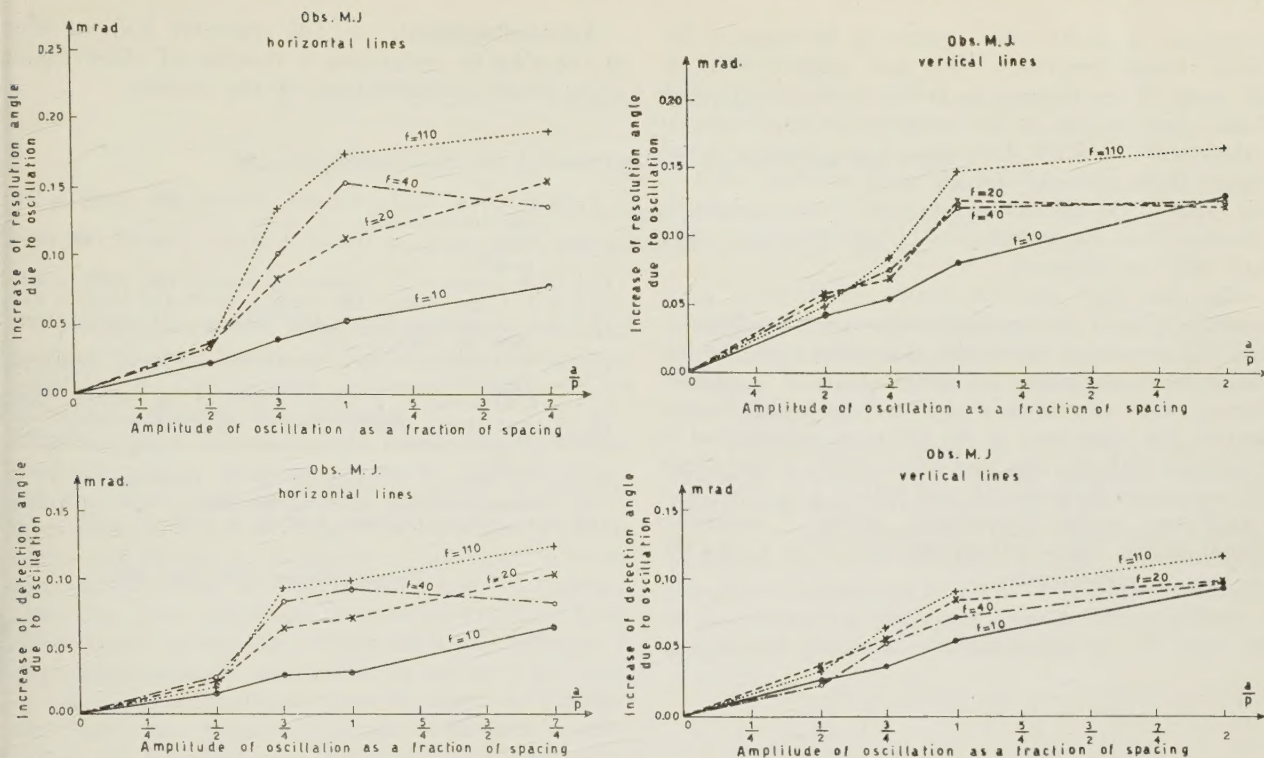


FIG. 8. — Increase of detection and resolution angles due to oscillation, as functions of amplitude, for (a) horizontal and (b) vertical grating. Observer M. J.

various curves correspond to measurements performed at different times.

These measurements are consequently affected by the unavoidable changes of the physiological conditions of the observer. The value of RP may vary even during the same observation. These facts affect in particular the values of DP and RP measured with the test object at rest. In some case DP or RP for an amplitude equal to 1 spacing are smaller than DP and RP for an amplitude equal to 2 spacings. The equivalent photometric profile has the same shape in both cases (fig. 2), but the contrast is smaller when the amplitude is twice the spacing: we should consequently expect that, at least for the highest frequencies, DP and RP would be smaller. This is not confirmed by the experiments, but we cannot rule out the possibility that our result is rather due to a variation of the physiological conditions of the observer than to a different effect of wider movements on vision. As regards in particular observer A. F. (fig. 3 *b*) it must be noticed that during observations with an amplitude of 1 spacing, this observer reported a great difficulty to detect and resolve vertical gratings.

For these reasons any conclusion derived from the experimental results must be accepted with some caution. However it seems that the following points can be established with a high degree of reliability.

1) Both DP and RP measured with an oscillating test object are smaller than those measured with the test object at rest. The decrease of DP is less pronounced than that of RP.

2) Both decreases are fairly regular when frequency increases from zero to 20-40 r. p. s.

3) In about 40 % of the curves the decrease of RP and DP seems to be followed by a little increase, when the frequency becomes a little greater than 20-40 r. p. s. A frequency of about 20-40 r. p. s. seems to cause a minimum of visual acuity.

4) When the frequency is greater than 50 r. p. s. both DP and RP oscillate about a value approximately constant. But several curves, especially those of observer A. M. E., show a general decrease also in the range of the highest frequencies.

5) The curves are more irregular in the range of high frequencies.

6) Curves corresponding to vertical test objects are more irregular than those corresponding to horizontal test objects.

Finally the curves of DP and RP as functions of amplitude of oscillation (fig. 6, 7, 8) suggest the following remarks.

7) For frequencies greater than 10 r. p. s. visual acuity always decreases when the amplitude varies from  $1/4$  to 1 spacing, whereas for greater amplitudes its behaviour becomes a little uncertain.

8) When the frequency is 10 r. p. s. visual acuity steadily decreases throughout the whole range of amplitudes.

It is evident from remarks 2 and 4 that the law of decreases of DP and RP undergoes a change about the

frequency of 40-50 r. p. s., which is the range of the flicker fusion frequency. This fact suggests that in the range of low frequencies both eye movements and flicker have a rôle in determining the experimental values of DP and RP. But, as said in remark 4, many curves show a steady though small increase, even in the range of the highest frequencies. It seems probable therefore that eye movements of high frequency may have also an influence.

The fact that the movements of 30-40 r. p. s. possibly give a minimum of visual acuity (observation 3) agrees with the results of COERMAN (11). He has investigated variations of visual acuity of a subject during rapid vibrations of the chair on which he was seated. The vibrations of the chair are transmitted to the subject's head, and they affect vision in the range of frequencies between 16 and 140 c. p. s. COERMAN found two critical frequencies, giving a minimum visual acuity. These lie between 25-40 c. p. s. and 60-90 c. p. s. respectively.

**Acknowledgement.** — The valuable help of Miss M. JEANNE in performing a number of observations is gratefully acknowledged by the authors.

#### REFERENCES

- [1] M. P. LORD, W. D. WRIGHT, *Nature*, **162**, 1948, p. 25 ; **163**, 1949, p. 803.
- [2] F. RATLIFF, L. A. RIGGS, *J. Exper. Psychol.*, **40**, 1950, p. 687.
- [3] H. B. BARLOW, *J. Physiol. (London)*, **116**, 1952, p. 290.
- [4] M. P. LORD, *Nature*, **169**, 1952, p. 1011 ; **170**, 1952, p. 670.
- [5] B. L. GINSBORG, *Nature*, **169**, 1952, p. 412. *Brit. J. Ophthalmol.*, **37**, 1953, p. 476.
- [6] R. W. DITCHBURN, B. L. GINSBORG, *J. Physiol. (London)*, **119**, 1953, p. 1.
- [7] G. C. HIGGINS, K. F. STULZ, *J. O. S. A.*, **43**, 1953, p. 1136.
- [8] F. RATLIFF, *J. Exper. Psychol.*, **43**, 1952, p. 163.
- [9] R. W. DITCHBURN, B. L. GINSBORG, *Nature*, **170**, 1952, p. 36.
- [10] L. A. RIGGS, F. RATLIFF, J. C. CORNSWEET, T. N. CORNSWEET, *J. O. S. A.*, **43**, 1953, p. 495.
- [11] R. COERMAN, *Luftfahrt Medizin*, **4**, 1939-40, p. 73.

*Manuscript reçu le 20 août 1955.*



## Lettre à l'éditeur

## Au sujet de l'aberration sphérique de l'œil

ALEXANDRE IVANOFF

Muséum National d'Histoire Naturelle, Paris.

La remarque de G. WESTHEIMER [1] concernant l'origine des coordonnées de mes courbes d'aberration sphérique de l'œil est très judicieuse. Par contre ses déductions me semblent exagérées.

Le fait que dans mes courbes l'aberration sphérique tend vers zéro en même temps que la distance à l'axe achromatique, implique en effet que les rayons infiniment voisins de l'axe le coupent au niveau de la rétine, autrement dit que la pointe de la caustique se forme juste sur la rétine. G. WESTHEIMER fait remarquer à juste titre que c'est là une supposition gratuite, étant donné que mes mesures se limitent à des distances à l'axe achromatique supérieures ou égales à 0,5 mm, et que par conséquent on ignore la valeur de l'aberration pour des distances à l'axe inférieures à cette valeur. Il remplace mon hypothèse par une autre, à savoir que la valeur de l'aberration sphérique reste pratiquement constante pour des distances à l'axe inférieures à 0,5 mm. Dans ces conditions, l'aberration sphérique de l'œil moyen serait sensiblement nulle et indépendante de l'état d'accommodation, et mes mesures correspondraient simplement à un défaut de mise au point variable avec l'effort d'accommodation, et autorisé par les conditions expérimentales (test constitué de deux fils verticaux, partie utilisée de la pupille réduite à deux petits disques).

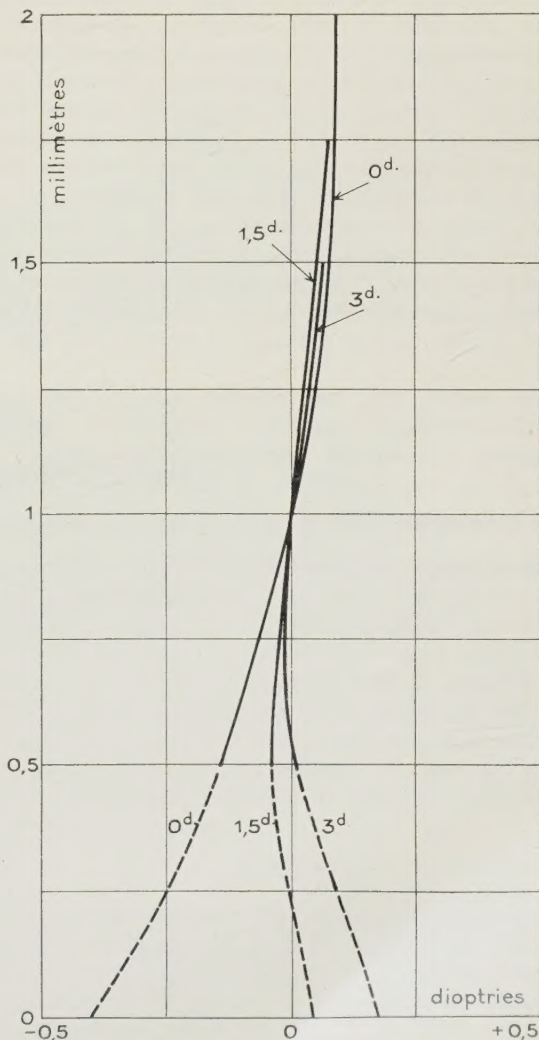
Je ferai remarquer tout d'abord qu'il est dangereux de raisonner sur une moyenne, surtout lorsqu'elle n'a pas grand sens, comme c'est le cas pour l'aberration sphérique de l'œil, très variable suivant l'œil étudié, et même suivant le méridien considéré [2]. D'ailleurs, s'il existe quelque raison de rapprocher l'aberration géométrique de l'œil de l'aberration dite sphérique [3], l'œil n'en est pas moins un système non centré, et son « aberration sphérique » n'est pratiquement jamais symétrique par rapport à l'axe achromatique. Si l'on examine séparément les résultats que j'ai obtenus sur chaque œil, et des deux côtés de l'axe, on verra qu'il est impossible de tout interpréter par un simple défaut de mise au point, et que l'hypothèse de G. WESTHEIMER (aberration constante pour des distances à l'axe inférieures à 0,5 mm) conduit souvent à des contradictions de part et d'autre de l'axe.

Par ailleurs, trois faits expérimentaux au moins semblent indiquer que l'aberration sphérique de certains yeux continue à varier pour des distances à l'axe inférieures à 0,5 mm :

1) En utilisant une toute petite région de la pupille, on a trouvé que le punctum remotum et le punctum

proximum d'un œil étudié varient avec la distance à l'axe, en sens contraires l'un de l'autre, et que cette variation continue nettement pour des distances à l'axe inférieures à 0,5 mm [4].

2) Parallèlement, lorsqu'on a obturé une partie axiale de la pupille de 1 mm de diamètre seulement, l'amplitude d'accommodation de l'œil étudié a diminué [5], ce qui indique que cette amplitude d'accommodation est particulièrement grande pour les rayons paraxiaux (le passage de l'œil de la sous-correction à la sur-correction, joint à la constriction pupillaire se produisant pendant l'accommodation, permet très





vraisemblablement à certains yeux d'économiser leur effort d'accommodation).

3) En diminuant le diamètre de la région utilisée de la pupille, on a trouvé sur deux yeux que l'hypermétropie continue d'augmenter rapidement pour des valeurs de ce diamètre inférieures à 1 mm [6].

Je pense donc que, malgré certaines observations qui tendraient à faire croire le contraire, l'aberration sphérique de l'œil varie souvent pour des distances à l'axe achromatique inférieures à 0,5 mm. Toutefois je concède très volontiers à G. WESTHEIMER que cette variation est exagérée sur mes courbes, l'aberration n'ayant pas obligatoirement une valeur nulle sur l'axe. Cette valeur peut d'ailleurs être obtenue par extrapolation d'après mes courbes, avec une précision pas trop mauvaise, puisque l'on connaît la valeur de l'aberration à 0,5 mm de part et d'autre de l'axe, ainsi que la pente générale de chacune des deux moitiés de la courbe.

La figure représente mes résultats moyens obtenus en 1953, en prenant égale à zéro l'aberration correspondant à une hauteur d'incidence égale à 1 mm, afin de ne pas faire jouer un rôle particulier à l'axe achromatique. J'ai tracé en trait discontinu les parties des courbes relatives à une distance à l'axe inférieure à 0,5 mm, obtenues par extrapolation comme il vient d'être dit ci-dessus, et par conséquent un peu incertaines (en fait, la variation de l'aberration au voisinage de l'axe semble être plus rapide pour certains yeux, mais serait sensiblement nulle pour d'autres).

Le passage de la sous-correction à la sur-correction, lorsque l'œil accommode de 3 d, se révèle même sur ces résultats moyens et même si l'on se limite aux distances à l'axe supérieures à 0,5 mm. Si l'on examine

séparément les résultats relatifs à chaque œil étudié, on observe que le phénomène est net pour 5 yeux, imprécis pour 4, et pratiquement inexistant pour un seul. Les contradictions entre les résultats obtenus par divers auteurs ne sont donc pas très étonnantes, d'autant moins que ma méthode fournit les résultats relatifs à un méridien, tandis que la méthode des pupilles annulaires fournit un résultat d'ensemble. La méthode de parallaxe que j'ai utilisée n'en reste pas moins la plus précise, le critérium d'alignement étant très sensible, l'axe achromatique pouvant être repéré avec précision, et les observateurs jouissant librement de leur vision binoculaire, ce qui assure la constance de la mise au point. L'imprécision relative aux rayons paraxiaux peut être supprimée en effectuant les mesures non plus par rapport à l'axe achromatique, mais par rapport à un rayon marginal de hauteur d'incidence connue.

Je remercie G. WESTHEIMER de sa critique très pertinente, qui m'a permis d'interpréter plus correctement mes résultats expérimentaux, et qui probablement mettra fin aux polémiques soulevées ces dernières années par l'aberration sphérique de l'œil.

#### RÉFÉRENCES

- [1] G. WESTHEIMER, *Opt. Acta*, 2, 1955, p. 151.
- [2] A. IVANOFF, *C. R. Ac. Sc.*, **231**, 1950, p. 373.
- [3] A. IVANOFF, *C. R. Ac. Sc.*, **231**, 1950, p. 526.
- [4] A. IVANOFF, *C. R. Ac. Sc.*, **230**, 1950, p. 1896.
- [5] A. IVANOFF, *C. R. Ac. Sc.*, **230**, 1950, p. 134.
- [6] A. ARNULF, O. DUPUY et F. FLAMANT, communication au colloque de Madrid, 1953.

*Manuscrit reçu le 20 décembre 1955.*

#### INFORMATION

##### Deutsche Gesellschaft für angewandte Optik

La réunion annuelle de la Société d'Optique allemande aura lieu à Berlin du 23 au 26 mai 1956.

Pour tous renseignements s'adresser au :

Dr. N. GÜNTHER, Finkenweg 2, Aalen/Württ.